B51

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C07K 13/00, C12N 15/18

A2

(11) International Publication Number: WO 94/24160

(43) International Publication Date: 27 October 1994 (27.10.94)

(21) International Application Number: PCT/US94/04361

(22) International Filing Date: 21 April 1994 (21.04.94)

(30) Priority Data:
08/049,802 21 April 1993 (21.04.93) US

(71) Applicant: BRIGHAM AND WOMEN'S HOSPITAL [US/US]; 75 Francis Street, Boston, MA 02115 (US).

(72) Inventors: BOISSEL, Jean-Paul, R.; 15 Tanager Street, Arlington, MA 02174 (US). BUNN, H., Franklin; 184 Islington Road, Auburndale, MA 02166 (US). WEN, Danyi; 9 Sherman Road, Chestnut Hill, MA 02167 (US). SHOWERS, Mark, O.; 110 Longwood Avenue, Brookline, MA 02146 (US).

(74) Agents: CIMBALA, Michele, A. et al.; Sterne, Kessler, Goldstein & Fox, Suite 600, 1100 New York Avenue, N.W., Washington, DC 20005-3934 (US).

(81) Designated States: AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KG, KP, KR, KZ, LK, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: ERYTHROPOIETIN MUTEINS WITH ENHANCED ACTIVITY

(57) Abstract

Novel modifications of erytropoietin (Epo) which improve the biological activity of the protein are provided. The modified Epo proteins (Epo muteins) may be used in the same manner as has been demonstrated for wild type Epo, except that relatively smaller doses are required due to the enhanced biological activity. Methods of using the Epo muteins for treatment of blood disorders are provided. A method of obtaining additional Epo muteins with enhanced biological activity is also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SI	Slovenia
SK	Slovakia
SN	Senegal
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
US	United States of America
UZ	Uzbekistan
VN	Viet Nam

Erythropoietin Muteins With Enhanced Activity

This invention was made with Government Support under grant nos. R01-HL42949 and R01-GM39900 sponsored by the National Institute of Health. The Government has certain rights in the invention.

Cross-Reference to Related Applications

This application is a continuation-in-part of U.S. Patent Application Serial No. 08/049,802 filed April 21, 1993 which is incorporated herein by reference.

Background

This invention relates to erythropoietin in general and more particularly to modified erythropoietin proteins (erythropoietin muteins) having improved biological activity.

Erythropoietin (hereafter referred to as Epo) is a glycoprotein hormone which, in its mature form in humans, is 166 amino acids long with a molecular weight of 34 to 38 kD (Jacobs *et al., Nature 313*: 806-809 (1985). Epo occurs in α, β, and asialo forms which differ slightly in their carbohydrate composition and biological activity (Dordal *et al., Endocrinology 116:* 2293-2299 (1985)).

In terms of biological function, Epo has been recognized as the hematopoietic cytokine that regulates the process of red blood cell production, known as erythropoiesis. Erythropoiesis is a controlled physiological process which normally produces red blood cells in numbers which do not impede blood flow, but which are sufficient for oxygen transport.

The binding of Epo to its cognate receptor (D'Andrea, A.D., *et al., Cell 57*:277-285 (1989)) on erythroid precursor cells in the bone marrow results in salvaging these cells from programmed cell death, known as

apoptosis (Koury, M.J., *et al., Science 248*:378-381 (1990)), allowing them to proliferate and differentiate into circulating erythrocytes (red blood cells). Epo thus facilitates the formation of erythrocytes from erythroid precursor cells.

The level of Epo in circulation normally controls the rate of red blood cell production in the body. Typically, Epo is present at a low plasma concentration sufficient to maintain a steady state concentration of red blood cells by stimulating formation of just enough red blood cells to replace those lost through normal processes, such as aging. However, when an increase in the production of red blood cells is needed, the amount of Epo in circulation is increased. This red blood cell deficit may be caused, for example, by anemia, the general loss of blood through hemorrhage, over-exposure to radiation, prolonged unconsciousness, or exposure to high altitudes where oxygen intake is reduced.

In mammals, Epo is produced in the fetal liver and adult kidney, circulates in the bloodstream, and binds to receptors on committed progenitor cells in the bone marrow and other hematopoietic tissues, resulting in proliferation and terminal maturation of erythroid cells (Jelkmann, W., *Physiol. Rev. 72*:449 (1992)). The expression of Epo, both mRNA and protein, is markedly increased by hypoxia, owing to a 3′ enhancer and to highly conserved elements in the promoter region (Semenza, G.L. *et al., Proc. Natl. Acad. Sci. USA 88*:5680-1684 (1988); Beck, I. *et al., J. Biol. Chem. 266*:15563-15566 (1991); Pugh, C.W. *et al., Proc. Natl. Acad. Sci. USA 88*:10553-10557; Blanchard, K.L. *et al., Mol. Cell Biol. 12*:5373-5385 (1992)). This elegant servomechanism enables Epo to regulate the red cell mass of man and other animals.

Genes encoding Epo from a number of species have been studied. Thus far, the Epo genes of man (Jacobs, K. *et al., Nature 313*:806-10 (1985); Lin, F.-K. *et al., Proc. Natl. Acad. Sci. USA 82*:7580-7584 (1985)), a monkey (*Macaca fascicularis*) (Lin, F.-K. *et al., Gene 44*:201-209 (1986)) and a rodent, the mouse (McDonald, J.D. *et al., Mol. & Cell Biol. 6*:842-848

(1986); Shoemaker, C.B. *et al.*, *Mol. & Cell Biol.* 6:849-858 (1986)) have been cloned, sequenced, and expressed. There is a high degree of sequence homology in the coding region of the mature secreted proteins from these species.

In particular, the gene encoding human Epo has been cloned and used to produce recombinant Epo in cell culture. Recombinantly produced human Epo has general utility as a substitute for Epo isolated from natural sources. In addition, recombinant Epo has facilitated the production of large amounts of Epo free from deleterious substances associated with naturally derived Epo, which has allowed for greater availability and utility of this protein.

Recombinantly produced Epo has proven especially useful for the treatment of patients suffering from impaired red blood cell production (Physicians Desk Reference (PDR), 1993 edition, pp 602-605). Recombinant Epo has proven effective in treating anemia associated with chronic renal failure and HIV-Infected individuals suffering from lowered endogenous Epo levels related to therapy with Zidovudine (AZT) (See PDR, 1993 edition, at page 602).

Modifications of the Epo protein which would improve its utility as a tool for diagnosis or treatment of blood disorders are certainly desirable. In particular, modified forms of Epo exhibiting enhanced biological activity would be more effective and efficient than native Epo in the therapy setting when it is necessary to administer Epo to the patient, enabling administration less frequently and/or at a lower dose. Administration of reduced amounts of Epo would also presumably reduce the risk of adverse effects associated with Epo treatment, such as hypertension, seizures, headaches, etc. (See PDR, 1993 edition, at pp. 603-604).

Unfortunately, available information regarding the structure of the Epo enzyme and its function does not permit the accurate prediction of modifications which would result in such improvements. In fact, the highly conserved nature of the amino acid sequence of Epo indicates that this protein is, for the most part, resistant to change and that most modifications would be

expected to have deleterious effects on Epo and its activity (See U.S. Patent No. 4,835,260 by Shoemaker).

Some speculation does, however, exist regarding the possibility of creating variants of Epo with altered characteristics that would be preferable to the wild type protein for particular purposes. A number of modifications of Epo have been proposed, including deletions, additions, and substitution of existing amino acids. These modifications have been proposed as a means of achieving a variety of effects, including separation of various Epo functions on individual protein fragments, improving the efficacy of recombinant Epo production, improving *in vivo* stability, etc. Of particular relevance to the present invention are those modifications proposed to improve the biological activity (i.e. ability to stimulate red blood cell production) of Epo.

U.S. Patent No. 4,703,008 by Lin, F-K. (hereinafter referred to as "the '008 patent") speculates about a wide variety of modifications of Epo, including addition, deletion, and substitution analogs of Epo. The '008 patent does not indicate that any of the suggested modifications would increase biological activity *per se*, although it is stated that deletion of glycosylation sites might increase the activity of Epo produced in yeast (See the '008 patent at column 37, lines 25-28). Also, the '008 patent speculates that Epo analogs which have one or more tyrosine residues replaced with phenylalanine may exhibit an increased or decreased receptor binding affinity.

Australian Patent Application No. AU-A-59145/90 by Fibi, M *et al.* (herein referred to as "Fibi") also discusses a number of modified Epo proteins (Epo muteins). Fibi generally speculates about the alteration of amino acids 10-55, 70-85, and 130-166 of Epo. In particular, additions of positively charged basic amino acids in the carboxyl terminal region are purported to increase the biological activity of Epo.

U.S. Patent No. 4,835,260 by Shoemaker, C. B. discusses modified Epo proteins with amino acid substitutions of the methionine at position 54 and asparagine at position 38. Such Epo muteins are thought to have improved

stability but are not proposed to exhibit any increase in biological activity relative to wild type Epo.

Beyond the few proposed modifications that have been speculated to enhance the biological activity of Epo, the art provides no further guidance regarding how Epo can be modified to increase its biological activity. This lack of guidance indicates that, based upon the present state of knowledge, other modifications of Epo which may increase biological activity cannot be predicted and may not even exist.

The present invention represents the inventors' achievement in overcoming the lack of guidance and unpredictability regarding modifications of the Epo protein which increase its biological activity.

Summary Of The Invention

It is therefore an object of the present invention to provide Epo muteins with enhanced biological activity.

It is a further object of the present invention to provide recombinant DNA encoding Epo muteins and methods of producing biologically active Epo muteins in a host cell culture using such recombinant DNA.

It is a further object of the present invention to provide methods of treating blood disorders involving abnormally low red blood cell populations using Epo muteins at dosages lower than that required for unmodified Epo to achieve the same affect.

It is a further object of the present invention to provide a method of obtaining additional Epo muteins with enhanced biological activity.

Brief Description Of The Drawings

Figure 1. **Oligonucleotide Primer Sequences.** Oligonucleotide primers corresponding to Epo DNA sequences are shown. The localization of the primer coincides with previously published nucleotide sequences for the human (Lin, F.-K. *et al.*, *Proc. Natl. Acad. Sci. USA* 82:7580-7584 (1985)) and murine (McDonald, J.D. *et al.*, *Mol. & Cell Biol.* 6:842-848 (1986)) Epo genes. EX2R and EX5 are not completely conserved between human and mouse (respectively 93% and 95% identity). An equal amount of each possible nucleotide was incorporated during the corresponding cycles of those primer syntheses. EX2R and NCO1 are reverse primers and their sequences represent the antisense DNA strand.

Figure 2. **PCR Strategy used for the cloning of mammalian cDNAs containing the complete coding sequence of the mature Erythropoietin Protein.** Genomic amplification and sequencing of exonic fragments, localized upstream and downstream from the nucleotide portion coding for the mature protein, allowed the design of species specific primers (SP). Utilization of those SP primers and/or of, sequences that are 100% conserved between man and mouse 5′ ATG and 3′ NCO1 primers (Figure 1) on cDNA templates prepared from kidney of uninduced or hypoxia-induced animals, allows the amplification of a large variety of mammalian Epo clones. Sense (→) and antisense (←) primers are represented by the arrows. Dashed boxes correspond to the coding part (propeptide and mature protein) of the five Epo exons. Gray boxes represent the 5′ and 3′ untranslated regions.

Figure 3. **Comparison of IV1/EX2R sequences from various mammals.** The numbering corresponds to the published human genomic sequence (Lin, F.-K. *et al.*, *Proc. Natl. Acad. Sci. USA* 82:7580-7584 (1985)). The reported human sequence was obtained from the amplification of purified genomic DNA from Hep3B cells and agreed with the sequence previously reported. The mouse sequence is from McDonald, J.D. *et al.*, *Mol. & Cell Biol.* 6:842-848 (1986). The horse sequence was obtained from

kidney-extracted genomic DNA. All the other sequences were established from PCR amplification of genomic DNAs purified from several mammalian renal-derived cell lines. The consensus sequence indicates the positions where a unique nucleotide was found in all the reported species. The boundary between intron 1 and exon 2 is represented by the ascending arrows. The locations of the two PCR primers, IVI and EX2R are shown by the dashed arrows.

Figure 4. **Alignment of the nucleotide sequences of mammalian Epo cDNAs.** The mouse sequence corresponds to the one previously reported (McDonald, J.D. *et al.*, *Mol. & Cell Biol. 6*:842-848 (1986)). PCR-produced human sequence was obtained from hypoxia-induced Hep3B cell line and is in total agreement with the previously published sequence (Jacobs, K. *et al.*, *Nature 313*:806-810 (1985); Lin, F.-K. *et al.*, *Proc. Natl. Acad. Sci. USA 82*:7580-7584 (1985)). Monkey (*Macaca mulatta*), rat, sheep, pig and cat were amplified from kidney-purified cDNAs. Amplifications of the human, monkey, rat and sheep were realized using the ATG (residues 1 to 20) and NCO1 (residues 723-742) primers. The line under residues 41 to 56 represents the specific SP1 primer used for the cloning of the pig and cat sequences. For these two mammals, 5' sequences for SP1 are derived from the data presented in Figure 3. Arrows indicate the start and the end of the coding sequence (propeptide and mature protein).

Figure 5. **Predicted Amino Acid Sequences of Epo Propeptide.** The amino acid sequences are derived from data obtained from both genomic IV1/EX2R and cDNA amplifications. The numbering corresponds to the human sequence. Ala 1 is the first amino acid of the human mature protein. The ascending arrows show the site of the cleavage by the signal peptidase, as determined for the human and cynomolgus monkey proteins.

Figure 6. **Alignment of the Primary Structures of mature mammalian Epo Proteins.** The human, Macaca fascicularis and mouse amino acid sequences were previously reported (Lin, F.-K. *et al.*, *Proc. Natl. Acad. Sci. USA 82*:7580-7584 (1985); Lin, F.-K. *et al.*, *Gene 44*:201-209

(1986); McDonald, J.D. *et al.*, *Mol. & Cell Biol.* 6:842-848 (1986)). The residue numbers correspond to the 166 aa of the mature human protein. Plain boxes indicate the positions of the predicted four α-helices in the human sequence (See Example II). Dashed boxes show the N- and O-glycosylation sites. Limits between exons are indicated by the vertical lines.

Figure 6A. **Schematic representation of the mammal erythropoietin.** The invariant amino acids among the eight sequences shown in Figure 6 are represented by the black boxes. The localization of each of the four a-helices is underlined. The three glycosylation sites are shown as diamonds. The main disulfide bridge is indicated by the heavy line connecting noncontiguous sequences. The small arrows under the sequence indicate the short SH-bridge, missing in the rodents.

Figures 7-7B. **Phylogenetic lineages derived from analyses of mammalian Epo cDNA sequences encoding the full length mature proteins from seven species.**

Figure 7. **Strength of groups in the maximum parsimony tree found on examining all 945 unrooted trees formed by seven terminal taxa.** This tree of lowest length required 374 base substitutions. Each link between ancestral nodes has a circled number; this strength of grouping number is the minimum number of substitutions that must be added to the length of the maximum parsimony tree to find a tree that breaks down the barrier (moves one or more sequences) between the two groups separated by the interior link.

Figure 7A. **The phylogenetic tree derived from the maximum parsimony reconstruction.** On the basis of other molecular evidence involving comparative amino acid sequence data from monotremes, marsupials, and many eutherian species (Czelusniak, J. et al., *Meth. Enzymol.* *183*:601-615 (1990); Czelusniak, J. *et al.*, In "Current Mammology", Volume 2, 3rd ed., H. H. Ganoways, pp. 545-572, Plenum Publ. Corp. (1990)) the root of this Epo phylogenetic tree is placed on the interior link separating the artiodactyl group from the primate, rodent, cat group. The numbers on the internodal links represent the numbers of base pairs by which

the nodal ancestral and descendant sequences differ. MYA = millions of years ago, as inferred from paleontological views of eutherian phylogeny.

Figure 7B. **Parsimony reconstruction on that portion of the cDNA sequences that are codons for amino acids.** The numbers shown as a fraction on each link are, in the numerator, the number of amino acid changing base replacements and, in the denominator, the number of silent base replacements. The computer algorithm that carried out this calculation is described in Czelusniak, J. *et al.*, *Nature* 298:297-300 (1982) (see the legend for Figure 2).

Figures 8-8A. **Phylogenetic lineages derived from analyses of the mammalian Epo intron 1-exon 2 sequences.**

Figure 8. **Strength of groups in the maximum parsimony tree found on examining all 135,135 trees formed by nine terminal taxa.** As the two feloid sequences (those of cat and lion) are nearly identical, they were treated as a single taxon in this strength of grouping analysis. The maximum parsimony tree for the segment (about 285 bp) of intron 1 and exon 2 from ten species required 361 base substitutions. The circled numbers are the strength of grouping numbers (as defined in the Figure 7 legend).

Figure 8A. **The near most parsimonious tree that groups dog with feloids.** This tree required 365 base substitutions. The numbers on links represent numbers of base pairs by which the nodal ancestral and descendant sequences differ.

Figures 9-9B. ***Model of the Three-Dimensional Structure of Erythropoietin.***

Figure 9. ***Ribbon diagram of Epo tertiary structure.*** The four α helices are labeled A to D; loops between helices are appropriately named. Disulfide bridges are shown and N- and O-glycosylation sites are indicated respectively by dark and dotted segments.

Figure 9A. Schematic representation of Epo's primary structure depicting predicted up-up-down-down orientation of the four antiparallel α

helices (boxes with arrowhead). This folding pattern is strongly suggested by the large size of the two interconnecting loops AB and CD. The limits of each helix were drawn according to Table II of Example I. A predicted short region of β-sheet is delineated by the dashed rectangle. The N-glycosylation sites are represented by the dotted diamonds, the O-glycosylation site by the dashed oval. The locations of the two disulfide bridges are shown as solid lines.

Figure 9B. Cross-section of the Epo molecule at the level of the four α helices. The helical wheel projections are viewed from the NH_2-end of each helix. The hydrophobic residues, localized inside the globular structure, are indicated by filled circles. The charged and neutral residues (open and gray circles respectively) are exposed at the surface of the molecule.

Figure 10. **Immunoprecipitations of wild type Epo and the Δ140-144 mutein.** Cos7 cells were transfected with pSG5, pSG5-Epo/wt or pSG5-Epo/Δ140-144. After three days, the cells were metabolically labeled with [^{35}S]-methionine and [^{35}S]-cysteine. Immunoprecipitations of cellular extracts and supernatants were performed with our polyclonal antibody, raised in rabbit against the native human Epo. The immunoprecipitates were analyzed by SDS-polyacrylamide gel electrophoresis. Lanes 2 to 4 correspond to cellular extracts, and lanes 5 to 7 correspond to culture supernatants, from Cos7 transformed with the following: plasmid without insert (lanes 2 and 5), wild type Epo (lanes 3 and 6), and Δ140-144 (lanes 4 and 7). Lane 1 represents the protein molecular weight standard. The two arrows show the normal secretion of the wild type Epo (35-37 kD) and the cytoplasmic retention of the mutein Δ140=144 (~28 kD).

Figures 11-11B. **Interconnecting Loop AB.**

Figure 11. **Schematic representation of the loop AB showing the localization of muteins with various deletions and amino acid replacements.** The dashed arrows point to the positions of the serine substitutions (in Δ48-52). The two N-glycosylation sites are represented by the gray diamonds. The small Cys29=Cys33 disulfide bridge is indicated.

Figure 11A. **Amount and biological activities of secreted muteins.** The upper bar graphs show the relative secretion of wild type and loop AB muteins as determined by radioimmunoassay. The lowest bar graphs display the calculated specific activity (ratio bioassay/RIA) for each mutein, in comparison with the value obtained for the wild type Epo (ratio = 100%).

Figure 11B. **HCD57 cell proliferation as a function of increasing concentration of wild type and serine-substituted Epo muteins.** HCD57 cells (10^4 per ml) were cultured for three days in a 96 well microtiter plate with media containing increasing concentrations of secreted proteins (mU=milli-units of Epo, a relative mass measurement). The line graphs show the cellular growth as measured by ^{3}H-thymidine uptake for cells cultured with wild type Epo (○), Epo mutein F48S (◇), Epo mutein Y49S (X), Epo mutein A50S (♦), Epo mutein W51S (□), and Epo mutein K52S (△). The number of viable cells was also measured with the MTT colorimetric assay and gave similar curves. Proliferation experiments using the human UT-7/Epo cell line (Komatsu, N., *et al., Cancer Res. 51*:341-348 (1991)) and the Krystal assay (Krystal, G., *Exp. Hematol. 11*:649-60 (1983)) produced identical results.

Figures 12-12A. **Interconnecting Loop CD.**

Figure 12. **Schematic representation of the loop CD showing the location of three deletion muteins: Δ105-109, Δ111-119, Δ122-126, and the insertion of seven residues after Lys116 (myc epitope).** The O-glycosylation site is indicated by the dashed oval.

Figure 12A. **Secretion and biological activities of the muteins located in loop CD.** The two bar graphs were created as described in

Figure 11A. The two mutants Δ111-119 and 116/myc were normally secreted and had full biological activities.

Figures 13-13A. **In Vitro Translation of the Epo Wild Type.**

Figure 13. **Analysis of the ^{35}S-labeled translation products by SDS-PAGE.** One-step transcription/translation reactions were performed in the SP6-TnT rabbit reticulocyte lysate system. 1/30 of each reaction was resolved on a 15% polyacrylamide gel. Lane 1-low Mr standard from Amersham; lane 2-*in vitro* reaction without added plasmid; lanes 3 and 4 are translation products obtained after incubation of 1 μg of circular p64T-Epo, in the presence or absence of canine pancreatic microsomal membranes, respectively.

Figure 13A. **Binding of the in vitro translated Epo wild type onto Epo Receptor-GTS-agarose beads.** $6x10^3$ cpm (counts per minute) of purified ^{35}S-labeled erythropoietin products, processed with microsomes (+) or not (-) were incubated in the presence of the extra-cytoplasmic domain of the Epo receptor (ERE_x), following the protocol described by Harris, K.W., *et al., J. Biol. Chem. 267*:15205-15209 (1992). Identical binding demonstrated that the conservation of the propeptide did not impair the hormone/receptor interaction.

Figures 14-14A. **COOH-end of Epo.**

Figure 14. **Schematic representation of the analyzed muteins,** corresponding to the deletion of the four last amino acids Δ163-166 and the replacements of the residues 162 to 166 by a KDEL or poly (His) sequences.

Figure 14A. **Relative secretion of these muteins.** The bioactivities in the supernatants (Krystal, G., *Exp. Hematol. 11*:649-60 (1983)) and the cell extracts (Komatsu, N., *et al., Cancer Res. 51*:341-348 (1991)) of transformed Cos7 cells were measured by *in vitro* proliferation assay using HCD57. More KDEL mutant remained in the cytosol of the Cos7, when compared with the wild type Epo and Δ163-166 or poly (His) muteins. However, all the analyzed muteins had the same specific activity as that of the wild type.

Figure 15. **Bacterial Expression of Wild Type Epo.**

Figure 15 Panel A. **Diagram of the fusion protein.** An NH_2-terminal 22 amino acid long peptide, containing a 10 histidine stretch, was fused to the mature erythropoietin sequence. Factor Xa cleavage allowed the recovery of the mature Epo with only two extra residues at its amino terminus.

Figure 15 Panel B. **IPTG induction of the fusion protein.** Transformed *E. coli* BL21 (DE3) cultures (OD_{600}=0.6) were grown in presence of 1 mM IPTG. Aliquots were collected at 0 (lane 2), 1 (lane 3), 2 (lane 4) and 3 hours (lane 5) and analyzed on a 15% SDS-polyacrylamide gel, stained with Coomassie Brilliant Blue. High level production of the fusion protein was rapidly obtained. Lane 6 corresponds to an aliquot from transformed bacteria grown for 3 hours in a medium without IPTG. Lane 1 is a low molecular weight standard.

Figure 15 Panel C. Purification of the fusion protein. After 3 hours of IPTG induction, the produced $(His)_{10}$-Epo was solubilized in 6 M guanidine-HCl and purified on a nickel affinity resin by increasing imidazole concentrations following the pET-His system protocol (Novagen). Samples of the column eluants were analyzed by SDS-PAGE and stained with Coomassie Brilliant Blue. Lane 1-elution by 20 mM imidazole; lane 2-elution by 100 mM imidazole, releasing the fusion protein; lane 3-chelation of the nickel by a 100 mM EDTA wash; lane 4-molecular weight standard.

Figure 15 Panel D. Detection of the E. coli recombinant Epo on a Western blot. Solubilized proteins were separated on a 15% SDS-polyacrylamide gel, transferred to a nitrocellulose membrane and probed with a 1/2000 dilution of our native wild type polyclonal antibody, as described under Materials and Methods. Lane 1-analysis after oxidative reduction; lane 2-after dialysis against the factor Xa buffer; and lane 3-after factor Xa cleavage.

Figure 16. **Relationship Between Production of Muteins and Proposed Secondary Structure.** This bar graph shows the amount of secreted proteins in the supernatants of transiently expressed Epo mutants, as

detected by radioimmunoassay. The muteins were aligned over a schematic representation of the native Epo molecule. Each deletion is shown as a stippled bar, the width of which is proportional to the number of residues deleted. The four α helices are represented by the black rectangles. The two disulfide bridges are indicated. These mutagenesis results are in good agreement with our proposed four α-helical model of Epo.

Figure 17. **Biological Activity of Helix A Muteins.** The top panel shows a helical wheel projection of Epo helix A with arrows indicating the positions where modifications were made to create muteins. The lower panel is a dose response curve representing the biological activity of wild type Epo (⊡) and helix A muteins E13A (♦), R14A (■), E18A (◇), K20A (✳), and E21A (□) over a range of dosages (X-axis; mU=milli-units of Epo, a relative mass measurement) as determined by the relative incorporation of [^{3}H]-thymidine measured in counts per minute (cpm) (Y-axis) into Epo dependent HCD57 cells.

Figure 18. ***Biological Activity of Helix D Muteins.*** The top panel shows a helical wheel projection of Epo helix D with arrows indicating the positions where modifications were made to create muteins. The lower panel is a dose response curve representing the biological activity over a range of dosages (X-axis) of wild type Epo (⊡) and helix D muteins D136A (♦), R139A (■), K140A (◇) and 143A (x) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into Epo dependent HCD57 cells.

Figure 19. **Biological Activity of Muteins Adjacent to Helix D.** The top panel shows the region adjacent to the C-terminal boundary of Epo helix D with arrows indicating the positions where modifications were made to create muteins. The lower panel is a dose response curve representing the biological activity over a range of dosages (X-axis) of wild type Epo (⊡) and muteins adjacent to helix D K152A (♦), K154A (■), and Y156A (◇) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into Epo dependent HCD57 cells.

Figure 20. **Comparison of Muteins With Increased Biological Activity.** A dose response curve representing the biological activity of wild type Epo (-*-), and muteins R143A (□) and K154A (♦) with increased biological activity relative to wild type Epo is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into Epo dependent HCD57 cells. n.

Figure 21. **Comparison of Muteins With Increased Biological Activity in Non-malignant Cells.** A dose response curve representing the biological activity of wild type Epo (-*-), and muteins R143A (□) and K154A (♦) with increased biological activity relative to wild type Epo, is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into murine spleen cells.

Figure 22. **Comparison of Muteins With Increased Biological Activity in a Human Cell Line.** A dose response curve representing the biological activity of wild type Epo (x), and muteins R143A (□), and K154A (♦) with increased biological activity (relative to wild type Epo) is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into human Epo-dependent UT-7/Epo cell line (Komatsu, N. *et al., Cancer Res. 51:* 341-348 (1991).

Figure 23. **Comparison of Muteins with Increased Biological Activity in a Human Cell Line.** A dose response curve representing the biological activity of wild-type Epo (⊡) and mutein N147A (♦) with increased biological activity (relative to wild-type Epo) is shown over a range of dosages (mU) (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into human Epo-dependent UT-7/Epo cell line.

Figure 24. **Comparison of Muteins with Increased Biological Activity in a Human Cell Line.** A dose response curve representing the biological activity of wild-type Epo (⊡) and mutein N147A (♦) with increased biological activity (relative to wild-type Epo) is shown over a range of dosages

(mU) (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into human Epo-dependent UT-7/Epo cell line.

Figure 25. **Comparison of Muteins with Increased Biological Activity in a Human Cell Line.** A dose response curve representing the biological activity of wild-type Epo (⊡) and mutein N147A (◆,■) with increased biological activity (relative to wild-type Epo) is shown over a range of dosages (mU) (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into human Epo-dependent UT-7/Epo cell line.

Biological Activities of Helix A Muteins.

Figure 26. **Comparison of Muteins with Increased Biological Activity in a Human Cell Line.** A dose response curve representing the biological activity of wild-type Epo (⊡) and mutein K20A with increased biological activity relative to the wild type Epo, is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into UT-7 Epo cells.

Figure 27. **Comparison of Muteins with Increased Biological Activity in a Human Cell Line.** A dose response curve representing the biological activity of wild-type Epo (⊡) and mutein Y49S with increased biological activity relative to the wild type Epo, is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into UT-7 Epo cells.

Biological Activities of Helix B Muteins.

Figure 28. **Comparison of Muteins with Increased Biological Activity in a Human Cell Line.** A dose response curve representing the biological activity of wild-type Epo (⊡) and mutein A73G with increased biological activity relative to the wild type Epo, is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into UT-7 Epo cells.

Biological Activities of Helix D Muteins.

Figure 29. **Comparison of Muteins with Increased Biological Activity in a Human Cell Line.** A dose response curve representing the

biological activity of wild-type Epo (⊡) and mutein K140A with increased biological activity relative to the wild type Epo, is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into UT-7 Epo cells.

Figure 30a. **Comparison of Muteins with Increased Biological Activity in a Human Cell Line.** A dose response curve representing the biological activity of wild-type Epo (⊡) and mutein R143A with increased biological activity relative to the wild type Epo, is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into UT-7 Epo cells.

Figure 30b. **Comparison of Muteins with Increased Biological Activity in a Mouse Cell Line.** A dose response curve representing the biological activity of wild-type Epo (⊡) and mutein R143A with increased biological activity relative to the wild type Epo, is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into HCD 57 cells.

Figure 31. **Comparison of Muteins with Increased Biological Activity in a Human Cell Line.** A dose response curve representing the biological activity of wild-type Epo (⊡) and mutein S146A with increased biological activity relative to the wild type Epo, is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into UT-7 Epo cells.

Figure 32. **Comparison of Muteins with Increased Biological Activity in a Human Cell Line.** A dose response curve representing the biological activity of wild-type Epo (⊡) and mutein N147A with increased biological activity relative to the wild type Epo, is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into UT-7 Epo cells.

Figure 33a. **Comparison of Muteins with Increased Biological Activity in a Human Cell Line.** A dose response curve representing the biological activity of wild-type Epo (⊡) and mutein K154A with increased

biological activity relative to the wild type Epo, is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into UT-7 Epo cells.

Figure 33b. **Comparison of Muteins with Increased Biological Activity in a Mouse Cell Line.** A dose response curve representing the biological activity of wild-type Epo (⊡) and mutein K154A with increased biological activity relative to the wild type Epo, is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into HCD 57 cells.

Figure 33c. **Comparison of Muteins with Increased Biological Activity in Non-Malignant Mouse Cells.** A dose response curve representing the biological activity of wild-type Epo (⊡) and mutein K154A with increased biological activity relative to the wild type Epo, is shown over a range of dosages (X-axis) as determined by the relative incorporation of [^{3}H]-thymidine (Y-axis) into primary murine spleen cells.

Figure 34. **Bioactivity of Epo Replacement Muteins.** The letters in the column on the left designate predicted helices (A, B. C, D) or interhelical loops (A-B, C-D). The 3'D designation (3' of D) in the table represents amino acids 153-166 which are at the carboxy-end of the D-helix. The amino acid of wild type human Epo, its position from the N-terminus, and its replacement according to the single amino acid letter code, are listed under the mutein column, e.g. S9A, represents a mutation from the wild-type human Epo at the amino acid serine of position number 9 to alanine. The three bioassay columns (human UT7 cell line, murine HCD57 cell line, and primary mouse spleen erythroid cells) show specific bioactivity of each mutein, expressed as a percentage of wild type human Epo bioactivity, with the background COS supernatant alone subtracted from the value. The mean and standard deviation of the bioassay values are listed for $n \geq 3$ determinations. The numbers within the brackets indicate the number of separate COS cell transfections over (/) the number of separate bioassays that were performed. NS: not secreted from COS 7 cells.

Definitions

In order to provide a clearer and more consistent understanding of the specification and claims, including the scope to be given such terms, the following definitions are provided.

Administration. The term "administration" is meant to include introduction of the Epo muteins of the invention into an animal or human by any appropriate means known to the medical or veterinary art, including, but not limited to, injection, oral, enteral, and parenteral (e.g., intravenous) administration.

Amino Acid Codes. The most common amino acids and their codes are described in the following table:

Table 1

Amino acid names and codes

Amino acid	Single letter code	Three letter code
Alanine	A	Ala
Arginine	R	Arg
Aspartic acid	D	Asp
Asparagine	N	Asn
Cysteine	C	Cys
Glutamic acid	E	Glu
Glutamine	Q	Gln
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophane	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

Animal. The term "animal" is meant to include all animals whose production of red blood cells (erythrocytes) is dependent upon, or stimulated by, erythropoietin (Epo). Foremost among such animals are humans; however, the invention is not intended to be so limiting, it being within the contemplation of the present invention to treat any and all animals which may experience the beneficial effect of the Epo muteins of the invention.

Efficacious Amount. An "efficacious amount" of an Epo mutein of the invention is an amount of such Epo mutein that is sufficient to bring about a desired result, particularly the stimulation of red blood cell production, especially upon administration to an animal or human.

Erythrocyte. The term "erythrocyte" is intended to refer to a red blood cell.

Erythroid Precursor Cells. By "erythroid precursor cells" is intended cells with a capability to proliferate and differentiate into erythrocytes that is dependent upon exposure or contact with Epo.

Erythropoietin (Epo). The term "erythropoietin", or "Epo" is primarily intended to refer to the mature form of this hormone. The terms "wild type Epo" and "native Epo" are used interchangeably to refer to Epo in its naturally occurring, unmodified form. Unless otherwise indicated, amino acid positions on the Epo protein referred to herein are made with reference to the mature, 166 amino acid human Epo sequence shown in Figure 6 and corresponding sequences from other species.

Host Cell. By "host" or "host cell" is intended the cell in which a gene encoding an Epo mutein of the invention is incorporated and expressed. An Epo mutein gene of the invention may be introduced into a host cell as part of a vector by transformation.

Mutein. By "mutein" is meant a mutant or modified protein with one or more modifications to its native amino acid sequence in the form of amino acid additions, deletions, or substitutions. An erythropoietin mutein refers to a modified erythropoietin protein.

Pharmaceutically Acceptable Vehicle. The term "pharmaceutically acceptable vehicle" is intended to include solvents, carriers, diluents, and the like, which are utilized as additives to preparations of the Epo muteins of the invention so as to provide a carrier or adjuvant for the administration of such Epo muteins.

Transformation. By "transformation" is intended the act of causing a host cell to contain a desired nucleic acid molecule, including either a native

Epo gene or a gene encoding an Epo mutein of the invention, not originally part of that cell using methods known in the art.

Transfection. By "transfection" is intended the introduction of a DNA or RNA vector carrying a desired nucleic acid molecule, including either a native Epo gene or a gene encoding an Epo mutein of the invention, into a host cell.

Vector. By "vector" is intended a DNA element used as a vehicle for cloning or expressing a desired gene, such as an Epo mutein gene of the invention, in a host.

Detailed Description of the Preferred Embodiments

This invention is drawn to modifications of Epo which enhance its biological activity. Due to their enhanced activity, the Epo muteins of the invention provide a preferable alternative to native Epo where it is used to stimulate red blood cell production (e.g. treatment of blood disorders, cell culture, etc.).

The present invention teaches a structural model of the Epo protein useful for identifying functional regions which may be modified to enhance biological activity. According to the invention, Epo is predicted to have a four anti-parallel amphipathic α-helical bundle structure. Based upon this predicted structure, functionally important regions predicted to reside on the external surfaces of the helices (designated A-D, see Figures 9A-9B) are identified. In particular, the external surfaces of helices A and D, predicted to be involved in Epo receptor binding, are identified. According to the invention, these functionally important regions serve as targets for modifications which may enhance biological activity. Increased activity relative to the wild-type Epo can be found with amino acid substitutions in the A,B,C or D helix.

According to one aspect of the invention, Epo muteins are provided in which one or more of the amino acids within the predicted external surfaces

of helices A, B, C, D and 3' of D, or within the region immediately adjacent to the C-terminal end of helix D, have been replaced. In particular, amino acid substitutions at positions 20, 49, 73, 140, 143, 146, 147 and 154 of wild type Epo have been replaced. The inventors have discovered that substitution of the amino acids normally occurring at these positions results in Epo muteins with significantly higher biological activity than wild type Epo proteins.

According to the invention, the amino acids at position 20 (typically alanine), 49 (typically serine), 73 (typically glycine), 140 (typically lysine), position 143 (typically arginine), position 146 (typically serine), position 147 (typically asparagine), and 154 (typically either lysine or threonine) of wild type Epo are preferably replaced with an alanine residue. In addition to alanine other amino acids with a chemical structure and properties similar to alanine, such as serine and threonine, are contemplated by the invention as suitable substitutes for achieving Epo muteins of the invention with enhanced biological activity.

In another aspect of the invention, a portion of the predicted interloop region between helices A and B extending from amino acid 48-52 is identified as important to Epo function. According to the invention, this functionally important region serves as another target for modifications which may enhance biological activity. The inventors have discovered that substitution of the amino acid at position 49 results in Epo muteins with significantly higher biological activity than wild type Epo proteins.

According to the invention, the amino acid at position 49 (typically tyrosine) of wild type Epo is preferably replaced with a serine residue. Other amino acids with a chemical structure and properties similar to serine, such as alanine and threonine, are contemplated as suitable substitutes for achieving Epo muteins of the invention with enhanced biological activity.

The modifications taught by the present invention occur in regions that are highly conserved among Epo proteins from evolutionarily divergent species (see Example I and Figure 6 in particular). These modifications are thus expected to be broadly applicable to Epo proteins which are substantially similar to the human Epo protein in the regions where the modifications occur,

including, but not limited to, Epo proteins from monkey, mouse, rat, sheep, pig, cat, and dog.

Biological activity of Epo muteins can be determined by assaying the ability of these proteins to stimulate the proliferation of Epo responsive cells such as murine spleen cells (Krystal, G., *Exp. Hematol. 11*:649-60 (1983); Goldberg, M.A., *et al., Proc. Natl. Acad. Sci. USA 84*:7972-7976 (1987)), murine Epo responsive murine erythroleukemia cells (Hankins, W.D., *et al., Blood 70*:173a (1987)), and human Epo-dependent UT-7/Epo cells (Komatsu, N., *et al., Cancer Res. 51*:341-348 (1991)).

When assaying for biological activity, it is important to recognize that, as taught by the present invention, high levels of biologically active Epo can cause terminal maturation and cessation of division of Epo responsive cells. Because of this possibly confounding phenomenon, it is important to determine biological activity based on a full dose response curve over a broad range of dosages which includes non-saturating levels of the particular Epo mutein being assayed.

Although the mechanism by which the modifications taught by the invention increase biological activity is not completely understood, one possibility is that these modifications enhance receptor binding affinity. This possibility is thought to apply particularly to the modifications at positions 20, 49, 73, 140, 143, 146, 147, and 154, which lie within a region predicted to be involved in Epo receptor binding according to the teachings of the present invention.

Genes Encoding Epo muteins (Epo mutein genes)

In addition to the Epo muteins themselves, genes encoding these muteins are also contemplated by the present invention. Genes encoding the Epo muteins of the invention may be made by modification of the native Epo gene using standard methods well known to those of skill in the art.

The native Epo gene may be obtained using a variety of standard techniques. For example, a DNA molecule corresponding to the known

sequence of the Epo gene from a number of species may be artificially synthesized (See Example I and Figure 4). The Epo gene may also be isolated from a cDNA or genomic library using nucleic acid probes based on available Epo DNA sequence information and standard hybridization techniques as taught in U.S. Pat. No. 4,703,008, issued Oct. 27, 1987, which is herein incorporated by reference (See also Jacobs *et al., Nature 313:* 806-810 (1985)). Alternatively, amplification of Epo gene sequences by polymerase chain reaction (PCR) may be accomplished using primers based on available sequence data as taught in Example I.

Once obtained, the wild type Epo coding sequence can be altered to encode an Epo mutein of the invention using standard oligonucleotide-directed *in vitro* mutagenesis techniques (See Zoller, M. and Smith, M., *Methods in Enzymology 100:* 468-500 (1983); Dente, L. *et al., Nucl. Acids Res. 11:* 1645 (1983); Kunkel, T.A., *et al., Meth. Enzymol. 154*:367-382 (1987); Vandeyar, M. *et al., Gene 65:* 129 (1988)). These techniques generally involve hybridization of an oligonucleotide carrying the desired alteration to a single stranded target DNA. This is followed by complementary strand synthesis of the target DNA which incorporates the oligonucleotide carrying the desired alteration. The complementary strand containing the desired alteration is then propagated in a bacterial host cell. Specific application of this type of technique to modify the Epo gene is described in Examples II and III and in Australian Patent Publication No. AU-A-59145/90 by Fibi *et al.* and in U.S. Pat. No. 4,835,260 by Shoemaker, both of which are herein incorporated by reference.

Expression of Epo mutein Genes

Genes encoding Epo muteins of the invention can be introduced and expressed in a selected host cell system using conventional materials and techniques. DNA elements such as promoters, enhancers, polyadenylation sites, transcription termination signals, and the like should be associated with

the Epo mutein coding sequence so as to promote and control expression of the Epo mutein. The specific regulatory elements used will depend upon the host cell system selected for expression, whether secretion of the protein is desired, and other considerations readily apparent to one of skill in the art.

Various vectors may be employed as vehicles for the introduction and expression of Epo mutein genes in a host cell. Such vectors useful in the different host cell types are well known and include, for example, the mammalian expression vectors pSG5 (Stratagene), p-RK1 (Genetics Institute), p-SVK3 (Pharmacia), p-EUK-C1 (Clontech), pCDM (Invitrogen), pc DNAI (Invitrogen), and the bacterial expression vectors pFLAG-1 (IBI), all pET system plasmids (Novagen), pTrcHis (Invitrogen), the pGEX series (Pharmacia), and pKK 233-2 (Clontech). These vectors may be maintained as episomes in the host cell or they may facilitate integration of the Epo mutein gene into the host cell genome, or both. Vectors may also include other useful features, such as genes which allow for the selection or detection of cells in which they have been successfully introduced.

Host cells suitable for expression of the Epo muteins of the invention include, but are not necessarily limited to, *E. coli*, yeast, insect cells, plant cells, and a variety of mammalian cell types, including in particular Chinese Hamster Ovary (CHO) cells, Cos7 cells, Cos1 cells, baby hamster kidney cells, and CV1 cells. Epo mutein genes can be introduced into a host cell using standard transformation or transfection techniques.

Host cells which provide for glycosylation are preferred for production of Epo muteins to be used *in vivo* since the carbohydrate structure of Epo is important for optimal biological activity *in vivo* (*See* Dordal, M.S. *et al.*, *Endocrinology 116(6):* 2293-2299 (1985)).

The Epo muteins of the invention may be recovered and purified from host cells in which they are expressed using known methods such as immunoaffinity chromatography with antibodies to human Epo.

Therapeutic Use of Epo muteins of the Invention

The Epo muteins of the invention are suitable for therapeutic use in animals, particularly humans, and may be used in the same manner as wild type Epo, except that lower dosages will be required to achieve the same level of biological activity.

Administration of the Epo muteins of the invention may be accomplished by any of the methods known to the skilled artisan, provided that the method effectively places the Epo mutein in the appropriate environment to exhibit its activity (i.e. in contact with erythroid precursor cells). A preferred method is by parenteral routes, including intravenous and subcutaneous administration.

Administration will ordinarily include an efficacious amount of the Epo mutein supplied in a pharmaceutically acceptable vehicle. The amount of Epo mutein which is efficacious will be determined by a trained professional depending upon a number of considerations, including the condition being treated and its severity, the sex and body weight of the subject being treated, the method of administration, and the presence of compositions in the pharmaceutically acceptable vehicle which affect Epo activity.

In any event, the efficacious amount of a Epo mutein as taught by the invention will be significantly less than the corresponding amount of unmodified Epo needed to achieve the same degree of efficacy. The difference between the efficacious amount of the Epo mutein and the analogous amount of unmodified Epo will correspond with the increased biological activity exhibited by the Epo mutein. It is contemplated that an efficacious amount of an Epo mutein of the invention will be as low as 10-fold less than native Epo, or as low as about 5-10 units/kg body weight for treating chronic renal failure patients or about 10-30 units/kg body weight for treating HIV-infected patients with serum Epo levels less than or equal to 500 mU/ml

(milliunits/milliliter) who are receiving a dose of AZT equal to, or less than, 4200 mg/week.

Since the Epo muteins of the invention can be used effectively at lower dosages relative to wild type Epo, their use may reduce potential adverse effects associated with administration of Epo such as exacerbation of hypertension, seizures, headaches, tachycardia, nausea, clotted vascular access, shortness of breath, hyperkalemia, and diarrhea (see PDR, 1993 edition, at pp. 603-604).

Further description of methods and materials useful in the recombinant production of the Epo muteins of the invention and their use is provided in the teachings of U.S. Patent No. 4,835,260 by Shoemaker and U.S. Patent No. 4,703,008 by Lin, both of which are herein incorporated by reference in their entirety.

The invention may be better understood and appreciated by reference to the following examples. These examples are provided merely for illustrative purposes and should not be considered in any way as limiting the scope of the claimed invention.

EXAMPLE I

Erythropoietin Structure-Function Relationships: High Degree of Sequence Homology Among Mammals

Summary

In order to investigate structure-function relationships of erythropoietin (Epo), we have obtained cDNA sequences that encode the mature Epo protein of a variety of mammals. A first set of primers, corresponding to conserved nucleotide sequences between mouse and human DNAs, allowed us to amplify by PCR intron 1-exon 2 fragments from genomic DNA of hamster, cat, lion, dog, horse, sheep, dolphin and pig. Sequencing of these fragments permitted the design of a second generation of species-specific primers. RNA was

prepared from anemic kidneys and reverse-transcribed. Using our battery of species-specific 5′ primers, we were able to successfully PCR amplify Epo cDNA from Rhesus monkey, rat, sheep, cat and pig. Deduced amino acid sequences of mature Epo proteins from these animals, in combination with known sequences for human, Cynomolgus monkey and mouse, showed a high degree of homology, which explains the biological and immunological cross reactivity that has been observed in a number- of species. Human Epo is 91% identical to monkey Epo, 85% to cat Epo and 80-82% to pig, sheep, mouse and rat Epos. There was full conservation of a) the disulfide bridge linking the NH_2 and COOH termini; b) N-glycosylation sites; and c) predicted amphipathic a-helices. In contrast, the short disulfide bridge (C29/C33 in human) is not invariant. Cys33 was replaced by a Pro in rodents. Most of the amino acid replacements were conservative. The C-terminal part of the loop between the C and D helices showed the most variation, with several amino acid substitutions, deletions and/or insertions. Calculations of maximum parsimony for intron 1/exon 2 sequences as well as coding sequences enabled the construction of cladograms that are in good agreement with known phylogenetic relationships.

INTRODUCTION

A large number of early physiologic studies have established extensive, though not complete, biological cross-reactivity between Epos of man and a number of other mammals including mouse, rat, sheep and rabbit (Jelkmann, W., *Physiol. Rev. 72*:449 (1992); Garcia, J.F. *et al.*, *Proc. Soc. Exp. Biol. Med. 112*:712-714 (1962); Zanjani, E.D. *et al.*, *Proc. Soc. Exp. Biol. Med. 131*:1095-1098 (1969); Biljanovic-Paunovic, L. *et al.*, *Period. Biol. 90*:421-427 (1988)). In contrast, non-mammalian vertebrates (amphibians (Rosse, W.F., *et al.*, *Blood 22*:66-72 (1963); Pinder, A. *et al.*, *J. Exp. Biol. 105*:205-213 (1983)), birds (Rosse, W. *et al.*, *Blood 27*:654-661 (1966); Black, C.P. *et al.*, *Respirat. Physiology 39*:217-239 (1980)), fish (Zanjani,

E.D., *et al.*, *Science 33*:573 (1969)) have erythropoietic hormones that fail to cross-react with mammalian erythroid cells, and vice-versa (Rosse, W.F., *et al.*, *Blood 22*:66-72 (1963); Pinder, A. *et al.*, *J. Exp. Biol. 105*:205-213 (1983); Rosse, W. *et al.*, *Blood 27*:654-661 (1966)).

Thus far, the Epo genes of man (Jacobs, K. *et al.*, *Nature 313*:806-10 (1985); Lin, F.-K. *et al.*, *Proc. Natl. Acad. Sci. USA 82*:7580-7584 (1985)), a monkey (*Macaca fascicularis*) (Lin, F.-K. *et al.*, *Gene 44*:201-209 (1986)) and a rodent, the mouse (McDonald, J.D. *et al.*, *Mol. & Cell Biol. 6*:842-848 (1986); Shoemaker, C.B. *et al.*, *Mol. & Cell Biol. 6*:849-858 (1986)) have been cloned, sequenced, and expressed. In view of the marked cross-reactivity between mammalian Epos, it is not surprising that there is a high degree of sequence homology in the coding region of the mature secreted proteins. In keeping with their close phylogenetic relationship, human and monkey Epos are 94% and 91% identical in nucleotide and amino acid sequence respectively. In contrast, human and mouse Epos are 76% identical in nucleotide sequence and 80% identical in amino acid sequence.

Direct information on the three-dimensional structure of Epo is not yet available. Insights into structure-function relationships of Epo can be gained from the analyses of a more complete set of animal sequences. Such information could be useful for sequence-based computer modeling of three-dimensional structure. Moreover, a larger data base would permit the identification of highly conserved domains that are likely to be crucial to folding and/or biological function. Finally, comparative amino acid and nucleotide sequence information on Epo provides additional data for investigating phylogenetic relationships.

In this Example, the use of PCR based techniques for obtaining the full coding sequences of Epos of Rhesus monkey, rat, sheep, pig and cat, as well as partial sequences from four other species is demonstrated.

Materials and Methods

Animal Samples. Human hepatoma Hep3B cell line and multiple kidney-derived cell lines from hamster (BHK), sheep (MDOK), pig (LLC-PK1), dog (MDCK), cat (CRF-K), lion (PAL1-K) and spotted dolphin (SP1-K) were obtained through the American Type Culture Collection (American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD 20852). Monolayer cells were grown in 100 x 20 mm tissue culture dishes using recommended media and maintained in a humidified 5% CO_2/95% air incubator at 37°C. In some experiments, cells were made hypoxic by overnight incubation in 1% O_2, 5% CO_2, 94% N_2 at 37°C.

Poly A+ RNA prepared from Rhesus monkey kidney was purchased from Clontech (Palo Alto, CA). Kidneys from anemic cat and horse were obtained from veterinarians at Tufts School of Veterinary Medicine and the University of Nebraska, respectively. Anemia was induced by three consecutive daily intraperitoneal injections of phenylhydrazine (60 mg/kg body weight) into male Sprague-Dawley rats (SD-strain) and by repeated bleeding of the sheep and the pig. Kidneys were aseptically removed after induction and stored immediately in liquid nitrogen.

A 4 kb genomic human Epo clone (g Epo4) was provided by Genetics Institute (Boston, MA).

DNA and RNA Preparations. DNAs from cell lines or homogenized kidneys were extracted using pancreatic RNAase/SDS/proteinase K following a procedure modified from Blin and Stafford (Blin, N. *et al.*, *Nucleic Acids Res. 3*:2303-2308 (1976)).

Kidneys were homogenized in 4 M guanidine isothiocyanate (10 ml/g of frozen organ), containing 0.1 M beta-mercaptoethanol and 10% N-lauryl-sarcosine. Total RNAs were isolated by centrifugation over 5.7 M CsCl (Chirgwin, J.M. *et al.*, *Biochemistry 18*:5294 (1979)). After ethanol precipitation, the samples were resuspended in diethyl pyrocarbonate-treated water and stored at -70°C. Confluent monolayer cells were washed twice with

sterile phosphate buffered saline and directly lysed in the 4 M guanidine isothiocyanate solution. Total RNAs were isolated as described above for the kidneys.

RNA samples were first converted into single strand cDNA. 2 to 4 μg of total RNA from kidney or 500 ng to 1 μg of total RNA from cultured cells were denatured at 68°C in the presence of 2 μg of oligo $dT_{(15)}$. Reverse transcription was carried out in a 20 μl final volume, containing 50 mM Tris-HCl pH 8.3, 6 mM $MgCl_2$, 40 mM KCl, 10 mM DTT, 500 μM dNTPs, 20 u RNAsin (Promega. Madison, WI) and 20 u of avian myeloblastosis virus reverse transcriptase (AMV-SuperRT, Molecular Genetic Resources. Tampa, FL). The reaction was allowed to proceed at 37°C for 20 minutes, then at 42°C for one hour. Inactivation of the enzyme was performed at 100°C for 10 minutes. cDNA samples were stored at -70°C.

Polymerase Chain Reaction. 20 to 100 ng of genomic-DNA or 0.5 to 1 μl of the RT-reaction were used as templates for amplification.

Standard PCR reactions were performed in a 100 μl volume containing 50 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% w/v gelatin, 200 μM dNTPs, 30 pM of each sense (5') and antisense (3') primers and 0.5 units of Taq DNA polymerase (Perkin Elmer Cetus, Emeryville, CA). After an initial denaturation step for 5 minutes at 95°C, routinely 35 cycles of PCR were performed on the DNA Thermal Cycler (Perkin Elmer Cetus). The denaturation step of each cycle was carried out at 94°C for 1 min. For each sample and/or the primer pair used, annealing temperatures were optimized (from 45°C to 60°C) for a 2-minute reaction. A 3-minute elongation step at 72°C ended each cycle.

Amplification products were analyzed on 1 to 3% agarose gel-TAE. Nu Sieve or Seaplaque agaroses (FMC., Rockland, ME) were used for preparative purposes. Specific PCR products were recovered from the gel, by the use of the Gene Clean II kit (Bio 101, La Jolla, CA), by Spin X centrifugation (Costar, Cambridge, MA) or by standard phenol/chloroform extractions of melted gel slices.

Subcloning and Sequencing. Small (250-330 bp) amplified genomic fragments were directly subjected to automated sequencing, using the PCR primers as DNA sequencing primers.

PCR-generated cDNA products were blunt-ended using Klenow fragment and 5' phosphorylated by T4 DNA kinase (Boehringer Mannheim, Indianapolis, IN). They were then cloned into the SmaI site of the phagemid pBluescript II SK (Stratagene, La Jolla, CA). The double-stranded plasmid was sequenced using the SK and KS sequencing primers.

In all cases, Sequencing reactions were carried out on an Applied Biosystems 373A Automated DNA Sequencer utilizing ABI's DyeDeoxy Terminator Sequencer kit (Foster City, CA) and thermal cycling with Taq DNA polymerase (Promega, Madison, WI) as previously described (Tracy, T.E. *et al.*, *Biotechniques 11*:68-74 (1991)). To avoid errors, for each species, samples derived from different amplification/cloning experiments were prepared and both strands were sequenced.

Mammalian Expression; Bioactivity. Vectors containing full-length mammalian Epo coding sequences were subcloned into pSG5 plasmid (Pharmacia, Piscataway, NJ) and were transiently expressed in Cos7 cells. 72-hour supernatants were tested for their ability to sustain cellular proliferation of the Epo-dependent HCD57 cell line (Sawyer, S.T. *et al.*, *Blood 72* (suppl. 1): 132a (1988)).

Results and Discussion

Because of the known biological cross-reactivity between various mammalian Epos and the presumption of strong homology in the coding sequences, a logical cloning strategy would be to screen cDNA libraries with moderate to low stringency. However, Epo mRNA is expressed at barely detectable levels in all tissues except in hypoxic kidney and liver (Goldberg, M.A. *et al.*, *Proc. Natl. Acad. Sci. USA 84*: 7972-7976 (1987); Fandrey, J. *et al.*, *Blood 81:* 617-623 (1993)). Therefore it would be necessary to

generate libraries for each of the species of interest. Our primary research interest is in structure-function relationships of Epo and therefore we require full coding sequences of the mature Epo protein, rather than full cDNA sequences. Accordingly, we elected to employ a PCR cloning strategy predicated on known strong homologies in Epo cDNA sequences that flank the region encoding the mature protein. There was sufficient conservation of sequence around the NCO1 site in the 3' untranslated region (UTR) that a single primer could be used to amplify full length mature Epo coding sequence in rat, sheep, pig and cat as well as dog and rabbit (data not shown). However, we were unsuccessful in finding a universal 5' primer and therefore had to generate species specific sequences from amplifications of genomic DNA. This two stage PCR strategy proved to be satisfactory in generating highly accurate sequence information with a minimum expenditure of time and resources. In all instances, there was full agreement between sequences of complementary strands.

Preparation of Mammalian Epo cDNAs

Primer Design for PCR Amplification of Epo. The sequences of Epo genes from three species have been already published: two primates, human (Jacobs, K. *et al.*, *Nature 313*:806-810 (1985); Lin, F.-K. *et al.*, *Proc. Natl. Acad. Sci. USA 82*:7580-7584 (1985)) and Cynomolgus monkey (*Macaca fascicularis*) (Lin, F.-K. *et al.*, *Gene 44*:201-209 (1986)) and a rodent, the mouse (*Mus muscuius*) (McDonald, J.D. *et al.*, *Mol. & Cell Biol. 6*:842-848 (1986); Shoemaker, C.B. *et al.*, *Mol. & Cell Biol. 6*:849-858 (1986)). There is substantial homology at the nucleotide level not only in the coding sequences, but also in portions of introns and in 5' and 3' untranslated sequences of exons 1 and 5, respectively. A number of primers, corresponding to conserved nucleotide sequences between mouse and-human, were synthesized and tested for their ability to produce genomic-PCR fragments from a wide variety of mammalian cell lines. A 4 kb human

genomic clone and genomic DNA extracted from the human hepatoma cell line Hep3B were used as control templates for the PCR amplifications.

In particular, two pairs of primers, IV1/EX2R and EX5/NCO1 (Figure 1), directed the amplifications of fragments of 330 and 250 bp (respectively) from genomic DNA of all human and mammalian cell lines that we investigated. IV1 is a 25-mer oligonucleotide localized in intron 1, 216/217 bases upstream from the start of exon 2. Human and mouse are identical except that the human sequence contains an extra G nucleotide at position 996. IV1 primer corresponding to the mouse sequence was synthesized. The 23-mer EX2R ends 28 nucleotides upstream from the 3′ end of exon 2. EX5 is a 22-mer, beginning 34 bases downstream from the 5′ end of exon 5. NCO1 represents a 20 bp, 100% conserved DNA fragment, starting 112/113 bases downstream from the TGA stop codon and contains the unique NCO1 site present in the human and mouse gene.

An ATG primer was also synthesized, corresponding to a 20 oligonucleotide stretch, extending both 5′ and 3′ from the initiator methionine codon. This sequence is also totally conserved between primates and mouse. While combinations of exonic primers (ATG1EX2R, EX5/NCO1 and ATG/NCO1) were able to direct amplification from cDNA prepared from RNA of the human Epo-producing cell line Hep3B, no amplification was obtained on cDNA prepared from uninduced or hypoxia-induced BHK, MDOK, LLC-PK1, MDCK, CRF-K, PAL1-K and SP1-K cells (results not shown). Therefore these renal cell lines apparently do not express Epo mRNA, either constitutively or after hypoxic stress, at a level detectable even by RT-PCR. Furthermore, no Epo protein was detectable by radioimmunoassay in the supernatants of cells maintained overnight in 1% O_2, 5% CO_2, 94% N_2 at 37°C (Goldberg, M.A. *et al.*, *Proc. Natl. Acad. Sci. USA* 84:7972-7976 (1987)).

The amplification of species specific IV1/EX2R and EX5/NCO1 genomic fragments allowed us to design a strategy shown in Figure 2 for obtaining cDNA clones containing the complete coding sequence of mature

Epo protein from various mammals. The IV1/EX2R primer pair resulted in a genomic fragment (~330 bp) corresponding to the 3′ third of the first intron and 80% of exon 2, encoding the COOH-end of the propeptide and the NH_2-terminal amino acids of the mature protein. EX5/NCO1 primers amplified a segment of exon 5, containing the COOH-end of the Epo protein and a 3′ untranslated sequence of about 140 bp downstream of the stop codon: DNA sequencing of these 5′ and 3′ fragments permitted, if necessary, the design of a second generation of species-specific primers (SP), localized outside the coding part of the mature protein (i.e., in the sequences encoding the propeptide and the 3′ untranslated region).

Amplification of Partial Intron 1/ Exon 2 Genomic Clones. We first explored the efficacy of the IV1 and EX2R primers for amplification of genomic fragments from different purified DNA templates. In all the reactions, fragments of predicted size were the main (and, when using the most stringent annealing temperature, often the only) products detected by agarose gel electrophoresis. The (direct) sequences of these PCR-generated fragments are presented in Figure 3. The nucleotide alignment includes 6 different orders of mammals: one primate, human (*Homo sapiens*); two artiodactyls, sheep (*Ovis aries*) and pig (*Sins scrofa*); one perissodactyl, horse (*Equus caballus*); one cetacean, spotted dolphin (*Stenella plagiodon*); three carnivores, dog (*Canis familiaris*), cat (*Felis catus*) and lion (*Panthera leo*); two rodents, mouse (*Mus musculus*) and hamster (*Cricetus cricetus*). Cat and lion exhibited an almost complete sequence identity, with only one T to G nucleotide substitution near the 5′ end of the amplified portion of the first intron.

The ability of the intronic IV1 to anneal to genomic DNA from various species and, in combination with EX2R, to amplify PCR fragments of identical size, demonstrated a high degree of conservation of sequence and position between mammals. This suggests that IV1 sequence may be involved to some extent in the regulation of the Epo gene expression. The comparison of IV1/EX2R sequences also revealed two remarkably conserved intronic

sequences: AT(T/A)GAATGAA(G/C)GC [SEQ. ID NO: 1] (nucleotides 1046 to 1058 in the human gene) and A(A/T)GGTN(G/C)GGG [SEQ. ID NO: 2] (nucleotides 1085 to 1094).

Amplification of Partial Exon 5 Fragments from Genomic DNAs. PCR reactions were also performed applying the EX5/NCO1 primer pair to various purified genomic DNAs. As previously observed for the amplification of IV1/EX2R, in all the tested samples a unique main band of about 250 bp (as predicted for human and mouse genes) was detected on analytical agarose gel electrophoresis. Direct DNA sequencing of these fragments demonstrated that the bulk of the purified PCR product corresponded to the expected Epo exon 5 sequence. However, for some DNA samples, purified from horse kidney and several cell lines (BHK and LLC-PK1 in particular), unknown sequences were present to various extents. These minor contaminating PCR products generated extraneous peaks in the sequence data, resulting in ambiguities at several positions. Nevertheless, computer-edited analyses of generated 3′ non-coding sequences were sufficient to design, if necessary, specific primers with greater than 95% accuracy (data not shown).

Cloning of Partial cDNAS Encoding the Complete Mature Epo Protein. As we previously mentioned, we were unable to obtain any amplification with cDNAs prepared from renal-derived mammalian cell lines. Therefore, we tested our battery of primers on reverse-transcribed RNAs prepared from kidney of several mammals. We have successfully amplified Epo cDNA from: one primate, Rhesus monkey (*Macaca mullata*), one carnivore, cat (*Felis catus*), one rodent, rat (*Rattus norvegicus*) and two artiodactyls, pig (*Sus scrofa*) and sheep (*Ovis aries*). The aligned nucleotide sequences are presented in Figure 3. They have been submitted to GenBank and have been assigned accession numbers.

Rhesus monkey, rat and sheep cDNAs were amplified using the ATG/NCO1 primer-pair combination. Pig and cat sequences were obtained by the use of a 5′ 24-mer specific primer (SP1-5′ TGCTTCTGCTATCTTTGCTGCTGC 3′) [SEQ. ID NO: 3]. IV1/EX2R

sequencing demonstrated that this exon 2 sequence was shared by the two species. In all PCR amplifications, NC01 was used as the reverse primer. Conservation of the NCO1 sequence among mammals suggests a potential biological role of this nucleotide sequence (perhaps in modulating the stability of Epo mRNA) (Rondon, I. J. *et al.*, *J. Biol. Chem.* *266*:16594-16598 (1991)). Considerable homology is also found in the sequenced 3' untranslated segment of exon 5 (Figure 4).

When compared with human Epo, the overall percentages of nucleotide identity are, respectively: 93.5 for the Rhesus monkey, 86.7 for the cat, 85.7 for the pig, 83 for the sheep and 75.5 for the rat. Rhesus monkey and Cynomolgus monkey (Lin, F.-K. *et al.*, *Gene* *44*:201-209 (1986)) were 99.6% identical. Mouse (McDonald, J.D. *et al.*, *Mol. & Cell Biol.* *6*:842-848 (1986); Shoemaker, C.B. *et al.*, *Mol. & Cell Biol.* *6*:849-858 (1986)) and rat nucleotide sequences showed greater than 93% identity. The two sequenced artiodactyls, pig and sheep, showed 88% identity.[1]

Comparison of Mammalian Epo Primary Sequences

Propeptides. The predicted amino acid propeptide sequences of several mammals are presented in Figure 5. In the primates, there was strong conservation of the sequences of deduced propeptides, with only two amino acid (aa) substitutions at positions -11 (Leu in humans vs. Val in both monkeys) and -2 (Leu in humans vs. Pro in both monkeys). Expression of Cynomolgus monkey Epo gene in cultured mammalian cells resulted in the production of mature monkey Epo that was elongated at the N-terminus by three additional residues: Val-Pro-Gly (Lin, F.-K. *et al.*, *Gene* *44*:201-209 (1986)). As the Rhesus monkey has the same substitutions, an identical site

[1] Several PCR attempts on total RNA purified from the horse's kidney were unsuccessful, even when using 5' and 3' specific equine primers (SP), presumably because analysis on 1% formaldehyde agarose gel demonstrated the degradation of the RNA.

of cleavage is likely. The Leu to Pro amino acid replacement at -2 is probably responsible for the differential activity of the signal peptidase observed between man and monkeys. The three rodents, mouse, rat, and hamster, were found to have identical propeptide amino acid sequences.

Mature protein. A comparison of the three previously published and our five deduced amino acid sequences of mature Epo proteins is shown in Figure 6. Erythropoietin is highly conserved among mammals. The amino acid alignment showed that more than 63% of the molecule is composed of invariant amino acids (106 residues) (Figure 6A). Most of the observed substitutions are conservative, involving residues with similar physical and chemical properties. The calculated percentages of sequence identity are shown in Table II below.

TABLE II
Degree of conservation among mammals of the mature Epo proteins

The percentages of identity among the various sequences were determined from the amino acid alignment reported in Figure 6.

		% identity
Primates	Man/Rhesus Monkey	90.5
	Rhesus/Cynomolgus	98.8
Rodents	man/mouse	79.8
	man/rat	82.1
	mouse/rat	93.5
Artiodactyls	man/sheep	81.0
	man/pig	81.6
	sheep/pig	88.7
Carnivores	man/cat	84.5

There is 100% conservation of Cys7 at the N-terminal portion of the polypeptide and Cys161 near the C-terminus. This disulfide bridge is essential

to the formation of a stable and functional cytokine (Wang, F.F. *et al.*, *Endocrinology 116*:2286-2292 (1985)). In contrast, while Cys29 is invariant in all the species we examined, Cys33 is present only in primates, sheep, pig and cat, but not in the two rodents where there is a proline. The lack of functional importance of this short disulfide loop is underscored by muteins in which tyrosine substitutions at each or both of these sites have no effect on Epo's biological activity (See Example II).

There is also 100% conservation of all three asparagine residues responsible for N-linked glycosylation. Even though Epo, deprived of its carbohydrate either by enzymatic cleavage (Takeuuchi, M. *et al.*, *J. Biol. Chem. 265*:12127-12130 (1990)) or by production in bacteria (Nahri, L.O. *et al.*, *J. Biol. Chem. 266*:23022-23026 (1991)) has full *in vitro* biological activity, survival of the hormone in the circulation depends upon N-linked glycosylation (Spivack, J.L. *et al.*, *Blood 73*:90-99 (1986)). In contrast, the O-linked glycosylation site (Ser126 in human Epo) is not essential for biological function since it is missing in mouse and rat Epos.

As shown in Figure 6, there is very high conservation of sequence in regions which, by algorithms based on primary structure, are predicted to be alpha helices (See Example II). The few differences in sequence within these helical regions are conservative replacements. Indeed, some sites within these predicted alpha helices are conserved in corresponding helices of other cytokines (Manavalan, P. *et al.*, *J. Prot. Chem. 11*:321-331 (1992)). In contrast, regions predicted to be inter-helical loops are less well conserved, particularly the 14 residue stretch between residues 116 and 130 of human Epo. In Example II a mutein with a deletion of residues 111 through 119 is shown to have normal stability and biological activity, whereas a deletion of residues 122 through 126 fails to produce a detectable protein, probably owing to markedly impaired stability.

Similar overall 4 α-helical bundle structure exists or is predicted for several other cytokines/hormones, such as growth hormone (GH), granulocyte-macrophage colony-stimulating factor (GM-CSF), IL-3, IL-4 and IL-5 (Bazan,

J.F. *et al., Immunol. Today 11:* 350-354 (1990)). Like Epo, GH exhibits a high degree of primary sequence conservation and biological crossreactivity between mammals. The mouse, rat, pig and sheep proteins show about 80% amino acid identity with human GH. In other cytokines, however, there is more sequence diversity among species. For example, the human and murine forms of IL-5, GM-CSF and IL-3 show respectively only 69, 67 and 26% amino acid homology. Furthermore, unlike Epo or GH, those proteins lack inter-species biological cross-reactivity.

Expression of the Mammalian Epo Hormones. Mammalian Epo cDNAs were subcloned into the pSG5 plasmid and transiently expressed in the monkey Cos7 cell line. Supernatants of the transfected cells were able to sustain the cellular proliferation of the murine Epo-dependent HCD57 cell line, demonstrating the biological cross-reactivity between species. As the human Epo antisera used in our radioimmunoassay bind to Epos from other species with variable affinity, the amount of produced protein was difficult to determine accurately. Experiments are in progress to further characterize the relative affinities of the various mammalian erythropoietins toward murine and human Epo receptors.

Phylogenetic Analyses of Epo Sequences. Nucleotide sequences encoding the full length mature Epo protein were analyzed by the maximum parsimony method (Czelusniak, J. *et al., Meth. Enzymol. 183*:601-615 (1990); Stanhope, M.J. *et al., Mol. Phylog. Evol. 1*:148-160 (1992)). An "all trees" set of computer programs determined the parsimony length of each of the 945 unrooted trees formed by the seven cDNA sequences (the terminal taxa) (Figure 4) and, on ordering those 945 trees according to increasing length, identified the minimum number of extra nucleotide substitutions needed to break up each group in the maximum parsimony (lowest length) tree. As shown in Figure 7 and 7A, the mouse and rat are strongly grouped, as are the human and monkey. The pig and sheep are more weakly grouped, in accord with accepted phylogenetic views, which place the divergence of these animals about 55-60 million years ago. As expected with the parsimony reconstruction

on only the codon sequences, silent base substitutions occurred much more readily than amino acid replacements (Figure 7B).

The Epo intron 1-exon 2 sequences shown in Figure 3 were also analyzed by the maximum parsimony method. Figure 8 shows strength of grouping results based on all 135,135 unrooted trees formed by nine terminal taxa (in this case, cat and lion sequences, which are nearly identical, were treated as a single taxon), and Figure 8A shows a phylogenetic tree that combines maximum parsimony with established phylogenetic evidence. As the most parsimonious tree (Figure 7) for these relatively short sequences failed to join dog to the feloid (cat and lion) branch (reflecting the fact that the ancestral split between canoids and feloids traces back to the early carnivores at about 60 million years ago), the phylogenetic tree shown in Figure 7A is a near most parsimonious tree in which dog joins the feloids. The rodents (mouse and hamster) are strongly grouped, as are the feloids (cat and lion), and at more moderate strength the cetacean (dolphin) is grouped with the artiodactyls (sheep and pig). There are other molecular data as well as some paleontological data that depict cetaceans originating from early artiodactyls (Czelusniak, J. *et al. Meth. Enzymol. 183*:601-615 (1990); Czelusniak, J. *et al.*, In "Current Mammology", Volume 2, 3rd ed., H. H. Ganoways, pp. 545-572, Plenum Publ. Corp. (1990); Irwin, D.M. *et al., J. Mol. Evol. 32*:128-144 (1991); Gingreich, P.D. *et al., Science 249*:154-157 (1990); Novacek, M.J., *Nature 356*:121-125 (1992)).

EXAMPLE II

Structure Function Relationships of Erythropoietin: Muteins that Test a Model of Tertiary Structure

Summary

On the basis of its primary sequence and the location of its disulfide bonds, we propose a structural model of the erythropoietic hormone erythropoietin (Epo) which predicts a 4 α-helical bundle motif, in common with other cytokines. In order to test this model, site directed mutants were prepared by high level transient expression in Cos7 cells and analyzed by a radio-immune assay and by bioassays utilizing mouse and human Epo dependent cell lines. Deletions of 5 to 8 residues within predicted α helices resulted in the failure of export of the mutant protein from the cell. In contrast, deletions at the N-terminus (Δ1-6), the C-terminus (Δ163-166), or in predicted interhelical loops (AB: Δ32-36, Δ53-57; BC: Δ78-82; CD: Δ111-119) resulted in the export of immunologically detectable Epo muteins that were biologically active. The mutein Δ48-52 could be readily detected by RIA but had markedly decreased biological activity. However, replacement of each of these deleted residues by serine resulted in Epo muteins with full biological activity. Replacement of Cys29 and Cys33 by tyrosine residues also resulted in the export of fully active Epo. Therefore this small disulfide loop is not critical to Epo's stability or function. The properties of the muteins that we tested are consistent with our proposed model of tertiary structure.

Introduction

Humoral regulation of red blood cell production was first proposed at the beginning of this century (Carnot, P., *et al., C.R. Seances Acad. Sci. 143*:432-35 (1906)). Convincing physiologic experiments documenting the existence of erythropoietin (Epo) (Krumdieck, N., *Proc. Soc. Exp. Biol. Med 54*:14-17 (1943); Reissmann, K.R., *Blood 5*:372-380 (1950); Erslev, A., *Blood 8*:349-357 (1953); Jacobson, L.O., *et al., Nature 179*:633-634 (1957)) were followed by its purification (Miyake, T., *et al., J. Biol. Chem. 252*:5558-5564 (1977)) and partial structural characterization (Lai, P.-H., *et al., J. Biol. Chem. 261*:3116-3121 (1986)). The molecular cloning of this biologically and clinically important cytokine (Jacobs, K., *et al., Nature 313*:806-810 (1985); Lin, F.-K., *et al., Proc. Nat. Acad. Sci. USA 82*:7580-7585 (1985)) has led to further understanding of its properties (Sasaki, H., *et al., J. Biol. Chem. 262*:12059-12076 (1987); Davis, J.M., *et al., Biochemistry 26*:2633-2638 (1987)).

The binding of Epo to its cognate receptor (D'Andrea, A.D., *et al., Cell 57*:277-285 (1989)) on erythroid progenitors in the bone marrow results in salvaging these cells from apoptosis (Koury, M.J., *et al., Science 248*:378-381 (1990)), allowing them to proliferate and differentiate into circulating erythrocytes. The Epo receptor is a member of an ever enlarging family of cytokine receptors (Bazan, F., *Biochem. Biophys. Res. Commun. 164*:788-796 (1989)). In like manner, Epo shares weak sequence homology with other members of a family of cytokines which also include growth hormone, prolactin, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, GCSF, GM-CSF, M-CSF, oncostatin M, leukemia inhibitory factor and ciliary neurotrophic factor (Parry, D.A.D., *et al., J. Mol. Recognition 1*:107-110 (1988); Bazan, F., *Immunology Today 11*:350-354 (1990); Manavalan, P., *et al., J. Protein Chem. 11*:321-331 (1992)). The genes encoding these proteins have similar numbers of exons as well as a clear relationship between intron-exon boundaries and predicted α-helical structure. These similarities

have led to the prediction that this family of cytokines share a common pattern of folding into a compact globular structure consisting of four amphipathic α-helical bundles. Such theoretical models of the structures of human growth hormone (Cohen, F.E., *et al., Proteins: Struct. Func. Genet. 2*:162-166 (1987)) and IL-4 (Curtis, B.M., *et al., Proteins: Struct. Func. Genet. 11*:111-119 (1991)) have been in remarkably good agreement with subsequent structures established by X-ray diffraction (HGH) (Abdel-Meguid, S.S., *et al., Proc. Natl. Acad. Sci. USA 84*:6434-6437 (1987); deVos, A.M., *et al., Science 255*:306-312 (1992)) or by multidimensional NMR (IL-4) (Redfield, C., *et al., Biochemistry 30*:11029-11035 (1991); Powers, R., *et al., Science 256*:1673-1677 (1992)). Moreover, the crystal structures of GM-CSF (Diederichs, K., *et al., Science 254*:1779-1782 (1991)) and monomeric M-CSF (Pandit, J., *et al., Science 258*:1358-1362 (1992)) are also in reasonable agreement with their predicted structures.

Thus far, the structure of Epo has not been analyzed by either X-ray diffraction or by NMR. In order to begin to gain an understanding of structure-function relationships, we have taken a three-pronged approach:

a) Sequence determination of Epo from mammals of different orders in order to establish regions of homology. This work is described in Example I.

b) Construction of a model of the three-dimensional structure of Epo, followed by the design and preparation of muteins that test this model. These experiments are presented in this Example.

c) Design and testing of muteins that provide information on receptor binding domain(s). This work is presented in Example III.

Materials and Methods

Computer-based Modeling of Structure

Prediction of Secondary Structure. Epo sequences from human, monkey, mouse, rat, sheep, pig, and cat were aligned (See Example I), and examined using a hierarchical approach to secondary structure prediction that assumes that these proteins are members of the α/α folding class (Levitt, M. *et al., Nature 261*:552-558 (1976)). First, the pattern-based method of Cohen, F.E. *et al., Biochemistry 25*:266-275 (1986) for turn prediction was used to delimit sequence blocks likely to contain secondary structure. Predictions using the methods of Garnier, J. *et al., J. Mol. Biol. 120*:97-120 (1978) and Chou, P.Y. *et al., Ann. Rev. Biochem. 47*:251-276 (1978) suggested α-helical regions within these blocks. Finally, helical wheel projections were used to examine and then limit helix length based on preserving amphipathic character as codified in the work of Presnell, S.R. *et al., Biochemistry 31*:983-993 (1992). The locations of glycosylation sites were also used to suggest helix boundaries.

Tertiary Structure Prediction. Earlier investigations have revealed the general principles of helix-to-helix packing in globular proteins (Richmond, T.J., *et al., J. Mol. Biol. 119*:537-55 (1978)). Exploring these principles, Cohen, F.E. *et al., J. Mol. Biol. 132*:275-288 (1979) developed a method for the generation of three-dimensional protein structures from the secondary structure assignment. These methods have been applied to myoglobin, tobacco mosaic virus coat protein, growth hormone, α- and β-interferon, IL-2,and IL-4 (Cohen, F.E., *et al., J. Mol. Biol. 132*:275-288 (1979); Cohen, F.E., *et al., Science 234*:349-352 (1986); Cohen, F.E., *et al., Proteins: Struc. Func. Gen. 2*:162-166 (1987); Sternberg, M.J.E., *et al., Int. J. Biological Macromolecules 4*:137-144 (1982)).

The algorithm for tertiary structure generation is divided into four computations. The program **aapatch** identifies clusters of hydrophobic residues within the putative helices that could mediate helix-helix interactions (Richmond, T.J., *et al., J. Mol. Biol. 119*:537-55 (1978)). **Aafold** generates all possible helix pairings according to the location and geometric preferences of the interaction sites. **Aabuild** generates the three-dimensional models of all possible structures from the list of helix pairings (from **aafold**) and subject to stearic restrictions and geometric constraints on chain folding. In the final step, **aavector** applies the user defined distance constraints (e.g., disulfide bridges) to the structures generated. At this stage, coordinates have been specified only for residues in the core α-helices. For residues in sequentially distinct loops, lower bounds on the inter-residue distances can be inferred from the relevant helix terminus.

Preparation of Epo Muteins

Construction of the Mutagenic/Mammalian Expression Plasmid. A M13 plasmid, containing a 1.4 Kb EcoRI-EcoRI human Epo cDNA insert (λHEpo FL12) was a gift from Genetics Institute (Cambridge, MA) (Jacobs, K., *et al., Nature 313*:806-810 (1985)). A 943 bp EcoRI-BglII fragment, corresponding to the complete coding sequence of the wild type human erythropoietin, including untranslated regions 216 bp upstream and 183 bp downstream, was inserted into the mammalian expression plasmid pSG5 (Stratagene) (Sambrook, J., *et al.*, In "Molecular Cloning: A Laboratory Manual", 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, pp. 1.38-1.41 (1989)) and named pSG5-Epo/WT.

Site-Directed Mutagenesis was carried out according to the protocol described by Kunkel, T.A., *et al., Meth. Enzymol. 154*:367-382 (1987). Single-stranded DNA was rescued from the pSG5-Epo/WT phagemid grown overnight in *E. coli* CJ236, in 2XYT media containing M13KO7 helper phage

(In Vitrogen) and 70 μg/ml kanamacyn (Sigma). The resulting uracil-containing ssDNA was used as a template for mutagenesis. Oligonucleotides (24 to 46 mer) were synthesized with their 5′ and 3′ ends complementary to the target wild type Epo sequence. A large variety of mutations (base substitutions, deletions and insertions) were created at the centers of the mutagenic primer sequences. Annealing of the phosphorylated primers (10:1 oligonucleotide/DNA template molecular ratio) was performed in 10 μl of a 20 mM Tris-HCl pH 7.4, 2 mM $MgCl_2$, 50 mM NaCl solution. The reactions were incubated at 80°C for 5 minutes, then allowed to cool slowly to room temperature over a one hour period. The DNA polymerization was initiated by the addition as a mix of 1 μl of 10X synthesis buffer (100 mM Tris-HCL pH 7.4, 50 mM $MgCl_2$, 10 mM ATP, 5 mM each dNTPs, 20 mM DTT), 0.5 μl (8 units) of T4 DNA ligase and 1 μl (1 unit) of T4 DNA polymerase (Boehringer Mannheim). After 2 hours at 37°C, 80 μl of 1X TE was added. 5 μl of the diluted reaction mix was used to transform competent *E. coli* NM522 (ung^+, dut^+).

Since a 40-80% mutation yield is normally obtained, 4 to 5 double-stranded plasmid clones from each reaction were sequenced with 7-deaza-dGTP and Sequenase (U.S. Biochemical) (Sanger, F., *et al., Proc. Natl. Acad. Sci. USA 74*:5463-5467 (1977)). As a rule, the entire coding sequences of the Epo mutants were examined for the presence of unwanted mutation by sequencing or restriction enzyme mapping.

Production of Wild Type Epo and Epo Muteins in Mammalian Cells. Cos7 cells grown to approx. 70% confluency were transfected with 10 μg recombinant plasmid DNA per 10 cm dish using the calcium phosphate precipitation protocol (Kingston, R.E., *et al.*, In "Current Protocols in Molecular Biology", Green Publishing Associates & Wiley Interscience, NY, pp. 9.1.1-9.1.9 (1991)). As a control of transfection efficiency, in several experiments 2 μg of pCH110 plasmid (Pharmacia) was co-transfected and β-galactosidase activity measured in the cytoplasmic extracts.

RNA Blot-Hybridization Analysis. Total RNAs were prepared from cultured Cos7 cells (Chirgwin, J.M., *et al., Biochemistry 18*:5294-5299 (1979)) and 2 μg samples electrophoresed on a 1.1% agarose gel containing 2.2 M formaldehyde. Transfer to GeneScreen Plus filters (New England Nuclear) and hybridization with a ^{32}P-labeled wt Epo probe were carried out as previously described (Faquin, W.C., *et al., Blood 79*:1987-1994 (1992)).

Quantitation of Transiently-Expressed Recombinant Epos. The amount of secreted protein in the supernatants of transfected Cos7 was determined by a radioimmune assay (RIA). The RIA was performed using a high titer rabbit polyclonal antiserum raised against the human wild type Epo and produced in our laboratory. ^{125}I-labeled recombinant Epo was obtained from Amersham. Details of the protocol have been published elsewhere (Faquin, W.C., *et al., Exp. Hematol. 21:* 420-426 (1993)).

Immunoprecipitation of ^{35}S-labeled Epo Proteins. Three days after transfection, the Cos7 monolayers were washed extensively with 1X PBS and the cells incubated for 20 minutes at 37°C in 2 ml of Met$^-$/Cys$^-$ Minimum Essential Media Eagle modified. In each culture dish, 100 μl of TRAN35S-LABEL ([^{35}S]-cysteine and [^{35}5]-methionine, ~ 10 mCi/ml, ICN Biochemical) was then added. After two hours, the conditioned media were harvested and cellular extracts were prepared by lysis in radioimmune precipitation buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.02% (w/v) sodium azide, 0.1% (w/v) SDS, 0.5% (w/v) sodium deoxycholate, 1% (w/v) Triton X-100, 1 mM phenylmethanesulfonyl fluoride and 1 μg/ml aprotinin). Samples were precleared with rabbit preimmune serum/protein A-Sepharose CL-4B (Pharmacia) for two hours. Immunoprecipitations were performed overnight with our polyclonal antibody specific for human recombinant wild type Epo and immunoadsorbed with protein A-Sepharose CL-4B. Immunoprecipitates were run on 15% SDS-polyacrylamide gels (Laemmli, U.K., *Nature 227*:680-685 (1970)) and analyzed by autoradiography after treatment with Enhance (New England Nuclear).

Bioassays. The dose-dependent proliferation activities of wild type Epo and Epo muteins were assayed *in vitro* using three different target cells: murine spleen cells, following a modification of the method of Krystal (Krystal, G., *Exp. Hematol. 11*:649-60 (1983); Goldberg, M.A., *et al., Proc. Natl. Acad. Sci. USA 84*:7972-7976 (1987)); Epo responsive murine erythroleukemia cell line, developed by Hankins (Hankins, W.D., *et al., Blood 70*:173a (1987)); and human Epo-dependent UT-7/Epo cell line, derived from the bone marrow of a patient with acute megakaryoblastic leukemia (Komatsu, N., *et al., Cancer Res. 51*:341-348 (1991)). After 22 to 72 hours of incubation with increasing amounts of recombinant proteins, cellular growth was determined by [^{3}H]-thymidine (New England Nuclear) uptake or by the colorimetric MTT assay (Sigma) (Yoshimura, A., *et al., Proc. Natl. Acad. Sci. USA 87*:4139-4143 (1990)).

Bacterial Expression. The wild type Epo target, corresponding to the nucleotide sequence coding for the mature protein, was PCR-amplified using appropriate primers. In the sense primer an NdeI site (CATATG) was placed immediately 5′ to the Ala1 codon of the mature protein. In the antisense primer a BglII site was placed 3′ to the TGA stop codon. After enzymatic digestion, the 516 bp PCR fragment was inserted in an NdeI/BamHI-cut pET16b plasmid (Novagen), which has a T7 promoter followed immediately by the lac operator. IPTG induction of transformed *E. coli* BL21 (DE3) (T7 RNA polymerase$^+$, Ion$^-$, ompT$^-$) resulted in high levels of expression of a fusion protein with a ten histidine stretch at the amino terminus. The oligo-His tag allowed the binding of the produced (His_{10})-Epo on a nickel affinity resin and its elution by increasing imidazole concentrations in presence of phenylmethylsulfonyl fluoride (Sigma). Most of the produced protein formed insoluble aggregates and was solubilized and affinity-purified under denaturing conditions in 6 M guanidine-HCl. Oxidative refolding was performed by overnight dialysis at 4°C against 50 mM Tris-HCl pH 8.0, 40 μM $CuSO_4$ and 2% (wt/vol) Sarkosyl. Soluble protein was further dialyzed against 20 mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM $CaCl_2$ and subjected to factor Xa

(New England Biolabs) cleavage to remove the NH_2-polyHis sequence. Monitoring of the fusion protein following induction and during the various steps of purification was done by electrophoresis on a 15% polyacrylamide-SDS gel, stained with Coomassie Brilliant Blue. Alternatively, the His-Epo fusion protein was detected on Western Blot (Gershoni, J.M., *et al., Anal. Biochem. 131*:1-15 (1983)), using a 1/2000 dilution of our wt native Epo polyclonal antibody and a second biotinylated rabbit-specific antibody which is detected with a streptavidin-alkaline phosphatase conjugate (Amersham).

In Vitro Transcription/Translation. Sense and antisense primers, creating new BglII sites respectively 5′ and 3′ of the initiator and stop codons, were used in a polymerase chain reaction on pSG5-Epo/WT template. After BglII cleavage, the 594 bp PCR-fragment was subcloned into p SP64T (Krieg, P.A., *et al., Nucl. Acid. Res. 12*:7057-7070 (1984)). This SP6 containing vector provides 5′ and 3′ flanking regions from Xenopus β-globin mRNA, which allow efficient *in vitro* transcription/translation. Previous experiments showed poor yields of *in vitro* translated protein, when using the GC-rich natural 5′ Epo untranslated region (UTR). One step *in vitro* transcription/ translation was carried out by incubation of 1 μg of circular p64T-Epo in a 50 μl reaction volume of SP6-TnT coupled rabbit reticulocyte lysate system (Promega), in presence of [^{35}5]-cysteine (1200 Ci/mM, New England Nuclear). In some cases, canine pancreatic microsomal membranes were added to the reaction mix. A purified GST-human Epo receptor extracellular domain fusion protein (EREx) was a gift from W. Harris and J. Winkelman, and the binding of the ^{35}S-labeled translation products onto EREx-glutathione agarose beads were performed as described (Harris, K.W., *et al., J. Biol. Chem. 267*:15205-15209 (1992)).

Results

Construction of a Model of the Three-Dimensional Structure of Erythropoietin

From an analysis of the putative Epo helix sequences, **aapatch** identified eight possible helix-helix interaction sites. In principle, these sites could be used to generate 1.6 x 10^4 structures. Of these, only 706 maintained the connectivity of the chain and were sterically sensible. These structures resembled four helix bundles, an increasingly common motif in protein structure (Presnell, S.R., *et al., Proc. Natl. Acad. Sci. 86*:6592-6596 (1989)). The structures that were not compatible with the native disulfide bridge between Cys_7 and Cys_{161} were eliminated. This reduced the total number of structures from 706 to 184 (total computer time approximately 1 hr. on a Silicon Graphics IRIS 4D/35G). The remaining structures were then rank ordered by solvent-accessible surface contact area, a measure of the validity of model structures. The most compact structures were right handed, all anti-parallel four-helix bundles with no overhand connections, but this may be an artifact of a failure to add the polypeptide chain that forms the loops to the helical core constructed by **aabuild**. The other less compact structures were left-handed four-helix bundles with two overhand loops, a topology previously seen in the structures of IL-4 and growth hormone. We suspect that this is the likely structure for Epo. The consensus for assignments of putative α-helices in human Epo are summarized in Table III. First, analysis of the topological distribution of known four-helix bundle structures indicate that nearly all examples have an antiparallel orientation (Presnell, S.R., *et al., Proc. Natl. Acad. Sci. 86*:6592-6596 (1989)). Second, the left-handed four-helix bundles with two overhand connections arranges the four amphipathic helices to form a compact hydrophobic core. Finally, a path for the loop regions of Epo can be imagined by analogy to IL-4 and growth hormone that preserves the compact globular nature of the Epo model structure. Figure 9 shows

schematic representations of the possible topological interactions between the four-anti-parallel α-helical bundles.

Table III

Predicted α-helical regions of the mature erythropoietin protein

Helix	N-terminus	C-terminus	Potential Helix-Helix Interaction Sites
A	9	22	19
B	59	76	63, 67, 70, 71
C	90	107	95, 102
D	132	152	141

Data were obtained using the various algorithms for secondary and tertiary structure generations described under Materials and Methods.

Several authors have suggested that the helical cytokines form a structural superfamily (Cohen, F.E., *et al., Science 234*:349-352 (1986); Parry, D.A.D., *et al., J. Mol. Recognition 1*:107-110 (1988); Bazan, J.F., *Biochem. Biophys. Res. Comm. 164*:788-795 (1989); Cosman, D., *et al., Trends Biochem. Sci. 15*:265-270 (1990); Bazan, J.F., *Proc. Natl. Acad. Sci. U.S.A. 87*:6934-6938 (1990)). On the basis of both the mature protein and the individual α-helices, Epo seems to be more closely related to growth hormone, prolactin, IL-6 and GM-CSF rather than the other members of the helical cytokine superfamily. Nevertheless, recent improvements in algorithms for the identification of distant evolutionary relationships between proteins from structural fingerprints suggested that it might be possible to align the IL-4 structure to the Epo sequences. The Eisenberg (Bowie, J.U., *et al., Science 253*:164-170 (1991)) structural environment and 3D-1D profile methods are a powerful tool for recognizing that a sequence is compatible with a known structure, e.g., a four-helix bundle. The NMR structure of IL-4 from Smith, L.J., *et al., J. Mol. Biol. 224*:899-904 (1992) was used to construct a 3D-1D profile. A mixture of sequences including four helix

bundles, globins, and non-helical structures were aligned against the IL-4 profile. Not surprisingly, the IL-4 structures from human and mouse gave the highest scores (Z^2 = 22.8 and 8.1). However, the other known four-helix bundle cytokines known to share a similar fold with IL-4, e.g., human growth hormone (Abdel-Meguid, S.S., *et al., Proc. Natl. Acad. Sci. 84*:6434-6437 (1987)) (Z = 2.3) and GM-CSF (Diederichs, K., *et al., Science 254*:1779-1782 (1991)) (Z = 2.3) fared no better than some globin sequences (Kuroda's and slug sea hare globin, Z = 5.0 and 4.8) that adopt a distinct tertiary structure. The results for the human and sheep Epo sequences were also ambiguous (Z = 1.6 and 0.8). These results suggest that while profile methods are a powerful tool for recognizing structural similarity, their failure to identify homology does not exclude the possibility that two proteins share a common fold. For distantly related or unrelated structures, current profile methods cannot replace *de novo* methods for tertiary structure prediction.

Design and Expression of Epo Muteins that Test the Proposed Structure

To test the proposed four α-helical bundle structure of erythropoietin and at the same time to attempt to locate functional domains, we created by site-directed mutagenesis a series of deletion, insertion and replacement mutants. These muteins were designed to analyze the principal predicted structural features of the molecule: α-helices, interconnecting loops, as well as the NH_2 and COOH termini. Structural and functional implications of the disulfide bridges and the glycosylation sites were also investigated.

α Helices. Short amino acid deletions were prepared in, or close to, the predicted A, B, C and D α helices. Human wild type and muteins were

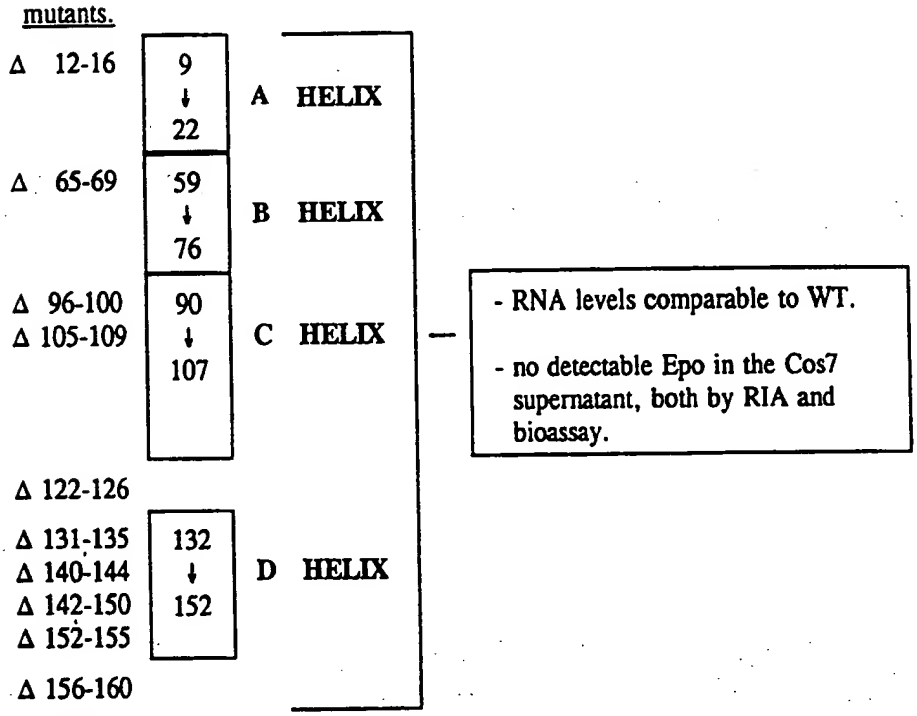
---

[2]Z-scores are used to describe the normalized weight associated with a profile score. A distribution is built from a collection of sequences with a mean 2 score of 0.0 and a standard deviation of 1.0. Z-scores greater than 6.0 are associated with significant alignments. Z-scores between 3.0 and 6.0 may or may not be structurally relevant.

transiently expressed in Cos7 cells. Northern blot analyses demonstrated that all the mutant plasmids produced about the same amount of mRNA as that of the wild type (data not shown). Yet, no detectable amount of Epo protein could be found in the Cos7 supernatants, either by radioimmunoassay or by bioassay using various Epo-dependent cell lines. Table IV summarizes these findings.

TABLE IV

Short deletions in, or close to, α helices



Predicted N and C termini of each α helix are indicated in the vertical boxes.

An example of SDS-PAGE of immunoprecipitants from *in vivo* ^{35}S-labeling is presented in Figure 10. As expected, when Cos7 cells were transfected with pSG5-Epo/WT, a 35-37 KD band was detected in the supernatant. In contrast, the deletion mutants (Table IV) could be detected in cellular extracts but were not exported from the cells. Figure 10 shows the

cytoplasmic retention of the mutein Δ140-144, lacking four residues in the middle of the predicted D helix. The apparent molecular weight (around 28 kD) is less than expected for a five-amino acid deletion. Therefore not only the secretion, but also the glycosylation, seem to be impaired. None of these muteins had deletion of glycosylation sites. It is likely that full glycosylation of Epo requires conservation of its molecular architecture. Similar results (reported in Table IV) were obtained for all the muteins having partial deletion of an α helical peptide segment.

Because contaminants in crude Cos7 cellular extracts severely interfere with the radioimmunoassay, no direct Epo quantitation was possible. However, aliquots of hypotonic extracts of Cos7 transfected with wild type Epo were able to sustain HDC57 proliferation. No similar biological activity was found for muteins with limited deletion of a helices.

Interconnecting Loops. The peptide segment joining A and B helices presents several interesting features (Figure 11). AB loop consists of 36 amino acids. Two N-glycosylation sites and a small disulfide bridge are located in the first half and their biological implications will be discussed later. The COOH end of the AB loop contains a stretch of amino acids that is strongly conserved among mammals (See Example I). Alignments of human, monkeys, cat, mouse, rat, pig, and sheep Epos showed a consensus sequence: DTKVNFYAWKR(M/I)(E/D)VG (residues 43 to 57) [SEQ. ID NO: 4]. Three deletions were constructed: Δ43-47, Δ48-52, and Δ53-57, and transiently expressed in Cos7. The amount of muteins detected by RIA in the supernatants of transfected cells was 10 to 40% lower than observed with wild type Epo (Figure 11A). Nevertheless, the three secreted muteins were biologically active. However, because Δ48-52 exhibited a marked decrease of the specific bioactivity, this site was studied in more detail by means of serine replacements. Krystal *ex vivo* bioassay as well as HCD57 and UT-7/Epo *in vitro* bioassays showed that these Ser mutants had biological activities similar to that of wild type (Figure 11B). Therefore, the observed decreases in both RIA and bioassay for the three deletion mutants are likely

to be the result of changes of structural conformation. The long length of loop AB may be critical for the up-up-down-down topography. A shorter AB segment may impose a strain on the interhelical connection. Chou and Fasman algorithms (Chou, P.Y., *et al., Ann. Rev. Biochem. 47*:251-276 (1978)) predicted a short β-sheet structure from residues 44 to 51 (<Pa>; <Pb>, 1.005<1.196). The presence of a short region of β-sheet in the connection between helices 1 and 2 (A and B) have been documented in the analyses of the three-dimensional structures of IL-4 (Redfield, C., *et al., Biochemistry 30*:11029-11045 (1991); Powers, R., *et al., Science 256*:1673-1677 (1992)), GM-CSF (Diederichs, K., *et al., Science 254*:1779-1782 (1991)), and monomeric M-CSF (Pandit, J., *et al., Science 258*:1358-1362 (1992)). In contrast, in human GH a short segment of a helix is found at the same location (Abdel-Meguid, S.S., *et al., Proc. Natl. Acad. Sci. USA 84*:6434-6437 (1987)). The structure/function implications of these short features are not yet understood.

Helix B is linked to helix C by a much shorter segment (residues 77 to 89) and contains in its center the third N-glycosylation site (Asn83). When the Δ78-82 mutein was expressed, a secreted protein was detected in the conditioned medium and conferred proliferative bioactivity on Epo-dependent cell lines (see Figure 16).

A similar long crossover connection (23 amino acids) is found between helix C and helix D. In contrast to what we previously observed for loop AB, a large deletion of nine residues at position 111-119 or a seven amino acid insertion of a myc epitope after residue 116 did not affect the secretion of these muteins (Figure 12). Furthermore, these two proteins had normal specific activity, as seen by the ratio of bioassay to RIA. Our rabbit polyclonal antibody raised against the native form of the human wild type fully recognized the two mutants, demonstrating that the overall spatial conformation of Epo was well preserved. According to the algorithm of Emini, E.A., *et al., J. Virol. 55*:836-839 (1985), the residues 111-119 are predicted to be at the surface of the molecule. Primary amino acid alignments

of mammalian Epo showed a large variation in the sequence of residues 116 to 130, including amino acid deletion, insertion, and substitution (See Example I). Surprisingly, when the deletion Δ122-126 mutein, which removed the O-glycosylation site (Ser 126), was transiently expressed in monkey cells, protein secretion was inhibited. Both rodents, rat and mouse, lack the O-glycosylation site because of a Ser126 to Pro replacement. Furthermore, when a Ser126 replacement mutein was expressed in normal CHO cells, (DeLorme, E., *et al., Biochemistry 331*:9871-9876 (1992)) or when wild type Epo was expressed in cells having a defect in O-linked glycosylation (Wasley, L.C., *et al., Blood 77*:2624-2632 (1991)), neither secretion nor biological activity were impaired. Therefore, failure of secretion of the Δ122-126 mutein may be the result of some other structural alteration. In particular, the proline residue at position 122 is invariant among mammals.

N and C Termini. Deletion of residues 2 to 6 only slightly affected the processing of a biologically active protein (see Figure 16). This deletion may impair cleavage of the propeptide, therefore explaining the lower yield of secreted Epo mutein in comparison to that of wild type. The fact that the mature monkey protein has an elongated (Val-Pro-Gly) N-terminus strongly suggests that the N-terminal part is not involved in the bioactivity of the molecule. Further evidence comes from the results, reported below, on the N-poly-His-Epo fusion protein expressed in *E. coli*, and also from the identical binding of *in vitro* translated ^{35}S-labeled wild type Epo onto EREX-glutathione agarose beads (Figure 13), with or without addition of canine pancreatic microsomal membranes which permit cleavage of the propeptide.[3]

The C-terminal sequence following helix D can clearly be divided into two distinct domains, separated by Cys161. The residues 151 to 161 were of special interest because they are highly conserved among mammals (See Example I). There are only two substitutions: Lys164 replaced by a Thr in

---

[3]All the mutants described in this example were subcloned into pSPG4T plasmid.

artiodactyls and cat, and Ala160 is replaced by a Val in mouse Epo. Both the Δ152-155 and the Δ156-160 muteins remained in the cytosol of the transfected Cos7 (Table IV). One possible explanation is that the residues 152 to 160 may, in fact, participate in the D helix. We predict that Gly151 is the break point of the structure. However, it is possible that this residue causes only a bend in the a helical structure and helix D may extend to Gly158.

The C-terminal part of the protein (residues 162 to 166) is clearly not involved in any structural or functional feature. Thus, the deletion of the four last amino acids or the replacement of residues 162-166 by either a KDEL sequence or a poly-histidine sequence[4] did not modify the specific activity of the erythropoietin (Figure 14). Radioimmunoassay revealed that the secretion of the KDEL-tail mutein in the media of transfected cells was 45% less than normally obtained with the wild type Epo. However, when compared to the wild type, this mutein had more biological activity in the hypotonic Cos7 cell extracts. The KDEL COOH-terminal sequence has been shown to be essential for the retention of several proteins in the lumen of the endoplasmic reticulum (Andres, D.A., *et al., J. Biol. Chem.* 266:14288-14282 (1991)). Nevertheless, because of overproduction in transiently expressed cells, a large percentage of recombinant protein escaped into the media.

Disulfide Bridges. Wang *et al.* demonstrated that the biological activity of Epo was lost irreversibly if the sulfhydryl groups were alkylated (Wang, F.F., *et al., Endocrinology 116*:2286-2292 (1985)). The mature human Epo has two internal disulfide bonds: Cys7=Cys161, linking the NH_2 and COOH termini of the protein and a small bridge between Cys29 and Cys33. Cys33 was previously changed to Pro by site-directed mutagenesis, and the resulting protein was reported to have greatly reduced *in vitro* biological activity (Lin, F.-K., *Molecular and Cellular Aspects of*

---

[4]The poly-His tail wild type mutant was purified by means of nickel affinity chromatography which will enable quantitation of cytosolic-retained mutants. The $(His)_6$ COOH-terminal sequence has been appended to all the muteins described in this example.

Erythropoietin and Erythropoiesis H8:23-36, NATO AS1 Series, Springer-Verlag, Berlin Heidelberg (1987)). However, rat and mouse Epos have the same substitution and yet exhibit full cross-species bioactivity. To resolve the role of the small disulfide bridge in human Epo function, we created a C29Y/C33Y double mutation. The resulting mutein was normally processed and showed the same *in vitro* bioactivity as the wild type (Figure 11A).[5] Furthermore, the deletion of five amino acid residues (Δ32-36) did not impair the secretion of a biologically active mutein. These data demonstrated that only the native and fully conserved disulfide bridge Cys7=Cys161 is crucial for the preservation of the molecular structure of erythropoietin.

Functional Role of the Glycosylation. Natural or recombinant human Epo is a heavily glycosylated protein; 40% of its molecular weight is sugars (Davis, J.M., *et al., Biochemistry 26*:2633-2638 (1987)). The protein has three N-linked oligosaccharide chains, located at amino acid positions 24 and 38 (in predicted loop AB) and position 83 (in loop BC). It has one O-linked carbohydrate chain at position 126 (in loop CD), which is missing in rodents. The role of these sugar chains in the biological activity of the human hormone has been extensively studied. Site-directed mutagenesis at the N-glycosylation sites demonstrated that even though the sugars were important for proper biosynthesis and secretion, their removal did not affect *in vitro* activity. This finding was corroborated by several investigators (Yamaguchi, K., *et al., J. Biol. Chem. 266*:20434-20439 (1991); Dubé, S., *et al., J. Biol. Chem. 263*:17516-17526 (1988)). However, Takeuchi, M., *et al., J. Biol. Chem. 265*:12127-12130 (1990) found that N-glycanase digestion results in almost complete loss of biological activity. In contrast, there is general consensus that glycosylation plays a key role in the biological activity of the hormone *in vivo*. Various reports have demonstrated that the N-linked sugar chains enhance the stability and survival of Epo in the blood stream (Fukuda,

---

[5]Two single replacement muteins (C29Y and C33Y) were also stable and had full biological activity (results not shown).

M.N., *et al.*, *Blood 73*:84-89 (1989); Tsuda, E., *et al.*, *Eur. J. Biochem. 188*:405-411 (1990)) and protect the hormone against clearance by the liver (Sasaki, H., *et al.*, *J. Biol. Chem. 262*:12059-12076 (1987)), thereby enabling the transit of the hormone from its site of production in the kidney to its target cells in the bone marrow (Spivak, J.L., *et al.*, *Blood 73*:90-99 (1989)).

We expressed the wild type Epo in *E. coli*. Accordingly, the produced protein completely lacks sugar. The pET expression system was used and is detailed in the Methods section. IPTG induction of transformed 8L21 (DE3) bacterial strain rapidly results in a high level of expression of the poly-His Epo fusion protein (Figure 15A and 15B). After three hours of induction, we obtained a typical yield of ~1 mg of total protein per ml of culture. However, the vast majority of the expressed protein was present in the inclusion bodies and therefore its solubilization and purification on the nickel beads was performed in 6 M guanidine-HCl. Oxidative refolding and factor Xa cleavage resulted in soluble forms (Figure 15C) and the *in vitro* biological activity was tested on HDC57 cells. The cleaved *E. coli* recombinant Epo showed a notable decrease of the specific activity (10% less than the fully glycosylated mammalian expressed protein), but was still able to maintain HCD57 proliferation. The observed reduction of *in vitro* activity is likely to be due to improper refolding of the insoluble protein and impaired physical stability of the *E. coli* Epo as previously reported (Narhi, L.O., *et al.*, *J. Biol. Chem. 266*:23022-23026 (1991)). However, the fact that the *E. coli* Epo was still able to trigger HCD57 growth indicated an overall preservation of the molecular structure. The uncleaved fusion protein, with 10 His residues at the N-terminus, exhibited a 67% loss in biological activity when compared to the cleaved protein. Thus, the addition of a 20-residue sequence to the N-terminus partially inhibited the biological activity.

Discussion

Currently the accrual rate of new protein sequences through gene cloning far outstrips the rate of determination of three dimensional structure. Epo is among a large number of biologically important proteins which have not yet been analyzed by X-ray diffraction or NMR. The problem is simplified by cumulative evidence that the structures of most proteins are likely to be variations on existing themes (Levitt, M., *et al., Nature 261*:552-558 (1976)). Indeed, as mentioned above, Epo appears to share common structural features with a large group of cytokines (Parry, D.A.D., *et al., J. Mol. Recognition 1*:107-110 (1988); Bazan, F., *Immunology Today 11*:350-354 (1990); Manavalan, P., *et al., J. Protein Chem. 11*:321-331 (1992)).

Computer-based prediction of structure can be reduced to a three stage process: secondary structure is predicted from the primary amino acid sequence and, when available, optical measurements. Analysis of Epo by circular dichroism reveals about 50% α helix and no detectable β sheet (Lai, P.-H., *et al., J. Biol. Chem. 261*:3116-3121 (1986); Davis, J.M., *et al., Biochemistry 26*:2633-2638 (1987)). With the knowledge of disulfide bonds, secondary structural elements are then packed into a set of alternative tertiary structures. The number of plausible arrangements can be reduced by empirical knowledge of preferred helix-helix packing geometries and the need for globular structure to form a hydrophobic core. The putative tertiary structure is then refined by standard force field calculations. Since there are a large number of alternate tertiary structures, the availability of experimentally determined structure of a homologous protein is critically important. Thus the predicted model of Epo structure gains considerable validity by knowledge of the structures of GH (Abdel-Meguid, S.S., *et al., Proc. Natl. Acad. Sci. USA 84*:6434-6437 (1987); deVos, A.M., *et al., Science 255*:306-312 (1992)) and IL*4* (Redfield, C., *et al., Biochemistry 30*:11029-11045 (1991); Powers, R., *et al., Science 256*:1673-1677 (1992)).

We have tested the predicted four anti-parallel α-helical bundle structure by means of site-directed mutagenesis. Deletions within predicted α helices would be expected to destabilize tertiary structure, whereas deletions or insertions in non-helical segments should be permitted unless they impose undue strain on the structure. For example, a deletion in an overhand inter-helical loop may result in insufficient length to connect the two helices. Results that we have obtained on muteins produced in mammalian (Cos7) cells are summarized in Figure 16. Our measurements of the quantities of processed mutein by RIA may underestimate the true amount of secreted Epo. Even a small deletion or insertion can result in a conformational change that may lead to impaired binding by our polyclonal antibody, raised against native human Epo. Thus the values for specific activity (biologic activity/RIA) that we report must be regarded as approximations. This caveat notwithstanding, our mutagenesis results are in good agreement with our proposed four α-helical model of erythropoietin. The proper folding of Epo into its native tertiary structure is necessary for stability and biological function. Muteins with short deletions inside predicted a-helices were not processed and exhibited no biological activity. In contrast, when deletions were created in predicted interconnecting loops, secreted proteins were detected, to varying degrees, both by radioimmunoassay and bioassay. Furthermore, additions or deletions at the N- or C-termini did not markedly impair the secretion and the biological activity of the Epo protein. Moreover, mutations at Cys29 and Cys33 showed that the small disulfide loop is not critical for biological activity. In order to delineate Epo's functionally important residues involved in the direct binding onto the Epo receptor, we have prepared and tested a series of amino acid replacements on the surfaces of the predicted α helices. These experiments are described in a Example III.

EXAMPLE III

Erythropoietin Structure-Function Relationships: Muteins That Test Functionally Important Domains

Introduction

The Epo receptor is a member of an ever enlarging family of cytokine receptors (Bazan, F., *Biochem. Biophys. Res. Commun. 164*:788-796 (1989)). In like manner, Epo shares weak sequence homology with other members of a family of cytokines which also include growth hormone, prolactin, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, G-CSF, GM-CSF, M-CSF, oncostatin M, leukemia inhibitory factor and ciliary neurotrophic factor (Parry, D.A.D. *et al.*, *J. Mol. Recognition 1*:107-110 (1988); Bazan F., *Immunology Today 11*:350-354 (1990); Manavalan, P. *et al.*, *J. Protein Chem. 11*:321-331 (1992)). The genes encoding these proteins have similar numbers of exons as well as a clear relationship between intron-exon boundaries and predicted α-helical bundles. Such theoretical models of the structures of human growth hormone (Cohen, F.E. *et al.*, *Proteins: Struct. Func. Genet. 2*:162-166 (1987)) and IL-4 (Curtis, B.M. *et al.*, *Proteins: Struct. Func. Genet. 11*:111-119 (1991)) have been in remarkably good agreement with subsequent structures established by X-ray diffraction (HGH) (Abdel-Meguid, S.S. *et al.*, *Proc. Natl. Acad. Sci. USA 84*:6434-6437 (1987); de Vos, A.M. *et al.*, *Science 255*:306-312 (1992)) or by multidimensional NMR (IL-4) (Redfield, C. *et al.*, *Biochemistry 30*:11029-11035 (1991); Powers, R. *et al.*, *Science 256*:1673-1677 (1992)). Moreover, the crystal structures of GM-CSF (Diederichs, K. *et al.*, *Science 254*:1779-1782 (1991)) and monomeric M-CSF (Pandit, J. *et al.*, *Science 258*:1358-1362 (1992)) are also in reasonable agreement with their predicted structures.

The structure of Epo has not yet been analyzed by either X-ray diffraction or by NMR. In order to begin to gain an understanding of structure-function relationships, we have taken a three-pronged approach:

a) Sequence determination of Epo from mammals of different orders in order to establish regions of homology. This work is described in Example I.

b) Construction of a model of the three-dimensional structure of Epo, followed by the design and preparation of muteins that test this model. These experiments are presented in Example II.

c) Design and testing of muteins that provide information on receptor binding domain(s). This work is presented in this Example.

MATERIALS AND METHODS

Construction of the Mutagenic/Mammalian Expression Plasmid. A M13 plasmid, containing a 1.4 Kb *Eco*RI-*Eco*RI human Epo cDNA insert (λHEpo FL12) was a gift from Genetics Institute (Cambridge, MA) (Jacobs, K. *et al.*, *Nature 313*:806-810 (1985)). A 943 bp *Eco*RI-*BglII* fragment, corresponding to the complete coding sequence of the wild-type human erythropoietin, including untranslated regions 216 bp upstream and 183 bp downstream, was inserted into the mammalian expression plasmid pSG5 (Stratagene) (Sambrook, J. *et al.*, In "Molecular Cloning: A Laboratory manual", 2nd ed. Cold Spring Harbor, Laboratory, Cold Spring Harbor, N.Y., vol. 1, pp.1.38-1.41 (1989)) and named pSG5-Epo/WT.

Site-Directed Mutagenesis was carried out according to the protocol described by Kunkel, T.A. *et al.*, *Meth. Enzymol. 154*:367-382 (1987). Single-stranded DNA was rescued from the pSG5-Epo/WT phagemid grown overnight in *E. coli* CJ236, in 2XYT media containing M13K07 helper phage (In Vitrogen) and 70 μg/ml kanamacyn (Sigma, St. Louis, Mo). The resulting uracil-containing ssDNA was used as a template for mutagenesis. Oligonucleotides (24 to 46 mer) were synthesized with their 5′ and 3′ ends complementary to the target wild-type Epo sequence. A large variety of mutations (base substitutions, deletions and insertions) were created at the

centers of the mutagenic primer sequences. Annealing of the phosphorylated primers (10:1 oligonucleotide/DNA template molecular ratio) was performed in 10 μl of a 20 mM Tris-HCL pH 7.4, 2 mM $MgCl_2$, 50 mM NaCl solution. The reactions were incubated at 80°C for 5 minutes, then allowed to cool slowly to room temperature over a one hour period. The DNA polymerization was initiated by the addition as a mix of 1 μl of 10X synthesis buffer (100 mM Tris-HCl pH 7.4 50, mM $MgCl_2$, 10 mM ATP, 5 mM each dNTPs, 20 mM DTT), 0.5 μl (8 units) of T4 DNA ligase and 1 μl (1 unit) of T4 DNA polymerase (Boehringer Mannheim). After 2 hours at 37°C, 80 μl of 1X TE was added. 5 μl of the diluted reaction mix was used to transform competent *E. coli* NM522 (ung^+, dut^+).

Since a 40-80% mutation yield is normally obtained, 4 to 5 double-stranded plasmid clones from each reaction were sequenced with 7-deaza-dGTP and Sequenase (U.S. Biochemical) (Sanger, F. *et al.*, *Proc. Natl. Acad. Sci. USA 74*:5463-5467 (1977)). As a rule, the entire coding sequences of the Epo mutants were examined for the presence of unwanted mutation by sequencing or restriction enzyme mapping.

Production of Wild-Type and Epo Muteins in Mammalian Cells. Cos7 cells grown to approx. 70% confluency were transfected with 10 μg recombinant plasmid DNA per 10 cm dish using the calcium phosphate precipitation protocol (Kingston, R.E. *et al.*, In "Current Protocols in Molecular Biology", Green Publ. Assoc. & Wiley Interscience, N.Y., vol. 1, pp. 9.1.1-9.1.9 (1991)). As a control of transfection efficiency, in several experiments 2 μg of pCH110 plasmid (Pharmacia) was co-transfected and β-galactosidase activity measured in the cytoplasmic extracts.

RNA Blot-Hybridization Analysis. Total RNAs were prepared from cultured Cos7 cells (Chirgwin, J.M. et al., *Biochemistry 18*:5294-5299 (1979)) and 2 μg samples electrophoresed on a 1.1% agarose gel containing 2.2 M formaldehyde. Transfer to GeneScreen Plus filters (New England Nuclear) and hybridization with a ^{32}P-labeled wt Epo probe were carried out as previously described (Faquin, W.C. *et al.*, *Blood 79*:1987-1994 (1992)).

Quantitation of Transiently-Expressed Recombinant Epos. The amount of secreted protein in the supernatants of transfected Cos7 was determined by a radioimmune assay (RIA). The RIA was performed using a high titer rabbit polyclonal antiserum raised against the human wild-type Epo and produced in our laboratory. ^{125}I-labeled recombinant Epo was obtained from Amersham. Details of the protocol have been published elsewhere (Faquin, W.C. *et al., Exp. Hematol. 21:* 420-426 (1993)).

Bioassays. The dose-dependent proliferation activities of wt and Epo muteins were assayed *in vitro* using three different target cells: murine spleen cells, following a modification of the method of Krystal (Krystal, G., *Exp. Hematol. 11*:649-660 (1983); Goldberg, M.A. *et al., Proc. Natl. Acad. Sci. USA 84*:7972-7976 (1987)); murine Epo responsive murine cell line, developed by Hankins (Hankins, W.D. *et al., Blood 70*:173a (1987)); and human Epo-dependent UT-7/Epo cell line, derived from the bone marrow of a patient with acute megakaryoblastic leukemia (Komatsu, N *et al., Cancer Res. 51*:341-348 (1991)). After 22 to 72 hours of incubation with increasing amounts of recombinant proteins, cellular growth was determined by [^{3}H]-thymidine (New England Nuclear) uptake or by the colorimetric MTT assay (Sigma) (Yoshimura, A. *et al., Proc. natl. Acad. Sci. USA 87*:4139-4143 (1990)).

RESULTS AND DISCUSSION

Our objective is to identify the functionally important domains on the Epo molecule by preparing muteins that have altered specific activity, i.e. significantly high or low biological activity per unit mass of protein. Such muteins must satisfy the following criteria: efficient secretion by Cos7 cells; full recognition (near normal binding affinity) by the polyclonal anti-Epo antiserum that we use in our radio-immune assay; preservation of the overall folding and three dimensional structure. In order to meet these criteria, the muteins that we employ in this study have only single amino acid

replacements. A comprehensive amino acid replacement scan of Epo would of necessity require the preparation and testing of a minimum of 166 muteins. Such an undertaking would be prohibitively labor intensive and costly.

In order to devise a rational strategy to reduce the number of required replacements, we have taken advantage of information we have obtained employing scanning deletion muteins (See Example II). On the basis of its primary amino acid sequence and disulfide bonds Epo, in common with other members of the cytokine family is predicted to have a four antiparallel amphipathic α-helical bundle structure. We have shown that deletions in non-helical regions at the N-terminus, the C-terminus and in loops between helices result in the formation of protein that is readily secreted from the cell and is biologically active. These regions can be ruled out as functionally important domains such as the sites involved in the binding of Epo to its receptor. These results suggest that the functionally important contact sites are likely to reside on the predicted α helices. Furthermore, since the amphipathic helices are predicted to bind to one another along their hydrophobic surfaces, the biologically relevant contact sites are likely to be residues predicted to be on the external surfaces of the helices. These considerations greatly restrict the number of mutations needed to provide a comprehensive and informative body of data.

Because structure-function studies on growth hormone and IL-4 have indicated that Helices A and D are involved in receptor binding, most of our muteins have been directed to these regions.

Results on Helix A muteins are shown in Figure 17. The top panel shows a helical wheel projection. Replacement by alanine of the predicted external residues 13, 18, 20 and 21 have no significant impact on biological activity as determined by the relative incorporation of [^{3}H]-thymidine into Epo dependent HCD57 cells. However, the replacement of Arg14 by Ala showed markedly impaired biological activity.

Figure 18, top panel, shows a helical wheel projection for Helix D. Nine of the nine predicted external residues were replaced by alanine.

Mutations at positions 136, 139 and 140 gave rise to proteins that were efficiently secreted and had full biological activity (Figure 18, bottom). In contrast Ala replacement of either Lys140, Arg143, Ser146, Asn147 or Lys154 resulted in muteins with significantly (3-fold) increased biological activity. Replacement of Asn147 with Ala resulted in the greatest increase in biological activity. (Figures 23-25).

Because the C-terminal boundary of Helix D is uncertain, we tested alanine replacements at positions 152, 154 and 156 which are predicted to be in the adjacent non-helical segment at the C-terminus. As shown in Figure 19, the Ala replacement at Lys154 had slightly increased biological activity while the Ala replacement at Tyr156 had slightly decreased biological activity.

The experiments reported above all involve testing the function of human Epo muteins on the mouse Epo responsive HCD57 cell line. Figure 20 shows a comparative display of those muteins that appear to have higher biological activity than wild-type Epo.

In order to assess the significance of these results we have also tested these muteins in a more physiologic system, mouse erythroid progenitors prepared from the spleens of animals challenged with phenylhydrazine. This bioassay obviates the possible confounding effects of a malignant clonal cell line such as HCD57 which was derived from mouse erythroleukemia virus.

As shown in Figure 21, results with erythroid cells from mouse spleen were very similar to those obtained with HCD57 cells.

In addition we have tested these muteins in an Epo responsive human cell line UT-7/Epo, derived from a patient with acute leukemia. This cell line provides the opportunity of examining the interaction of human muteins with the human Epo receptor. As shown of ordinary in Figure 22, these results closely parallel those with the two mouse cell bioassays.

It should be noted that in several of the dose response curves, at the highest concentrations of Epo tested, there is a significant drop-off in the incorporation of [^{3}H]-thymidine. This is due to the fact that high levels of biologically active Epo causes terminal maturation and cessation of division.

Because of this possibly confounding phenomenon, it is important to have, as in the experiments that we report here, a full dose response curve for each mutein with an adequate number of data points.

Taken together our results show that a restricted number of residues on Helix A and on Helix D are involved in Epo's biological activity. Further experiments are in progress to determine if the observed alterations in biological activity can be explained by parallel changes in the binding affinity to the Epo receptor. In addition, at the sites that we have identified as involved in Epo's biological activity, we plan to make a set of additional replacement muteins in order to further refine the ligand-receptor interaction. We are particularly interested in developing muteins with maximally increased biological activity, e.g. by preparing compound or multiplex replacements, each of which confers optimal enhancement of biological activity. Such a "superEpo" could have improved utility in the therapy of anemias.

Conclusions

We have produced a number of erythropoietin muteins comprising native erythropoietin with one or more modifications which improves the biological activity.

All references mentioned herein are incorporated by reference into this disclosure.

Having now fully described the invention by way of illustration and example for purposes of clarity and understanding, it will be apparent to those of ordinary skill in the art that certain changes and modifications may be practiced within the scope of the invention, as limited only by the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Brigham and Women's Hospital
75 Francis Street
Boston, MA 02115

INVENTORS: Boissel, Jean-Paul R.
Bunn, H. Franklin
Wen, Danyi
Showers, Mark O.

(ii) TITLE OF INVENTION: Erythropoietin Muteins With Enhanced Activity

(iii) NUMBER OF SEQUENCES: 59

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Sterne, Kessler, Goldstein & Fox
(B) STREET: 1100 New York Avenue, Suite 600
(C) CITY: Washington
(D) STATE: D.C.
(E) COUNTRY: U.S.A.
(F) ZIP: 20005-3934

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: (to be assigned)
(B) FILING DATE: Herewith
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/049,802
(B) FILING DATE: 21-APR-1993
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Cimbala, Michele A.
(B) REGISTRATION NUMBER: 33,851
(C) REFERENCE/DOCKET NUMBER: 0627.336PC01

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (202) 371-2600
(B) TELEFAX: (202) 371-2540

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATWGAATGAA SGC

13

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AWGGTNSGGG 10

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TGCTTCTGCT ATCTTTGCTG CTGC 24

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 12..13
(D) OTHER INFORMATION: /label= Peptide
/note= "The amino acid at position 12 may also be isoleucine. The amino acid at position 13 may also be aspartic acid."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TGAAGTTTGG CCGAGAAGTG GATGC 25

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AAGAKGTACC TCTCCAGRAC TCG 23

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CTGCTCCACT CCGAACAMTC AC 22

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTGGAGTGTC CATGGGACAG 20

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGGCGCGGAG ATGGGGGTGC 20

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 286 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TGGTAGCTGG GGGTGGGGTG TGCACACGGC AGCAGGATTG AATGAAGGCC AGGGAGGCAG 60

CACCTGAGTG CTTGCATGGT TGGGGACAGG AAGGACGAGC TGGGGCAGAG ACGTGGGATG 120

AAGGAAGCTG TCCTTCCACA GGCCACCCTT CTCCCTCCCC GCCTGACTCT CAGCCTGGCT 180

ATCTGTTCTA GAATGTCCTG CCTGGCTGTG GCTTCTCCTG TCCCTGCTGT CGCTCCCTCT 240

GGGCCTCCCA GTCCTGGGCG CCCCACCACG CCTCATCTGT GACAGC 286

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 290 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CTCGTAGTTG TTGGGTGGCG TGTGCACGCG GTCAGTGTAT TGAATGAAGG CAAAGGAGGC 60

AGAGCCTGCG CGCTCGCAAG GTTGGGGTCG GAAACGGCTG GCTGGGGGCA AGACGGCCGG 120

ATGGGTGAAC CTGTCCGTCT AAACCAACCC TTCTCCTGCA ATGCCCAGCC TCACATTCAG 180

CCTGGCTCTC TTTCCTAGAA TGTCCTGCCC GGCTGCTTCT GCTATCTTTG CTGCTGCTTC 240

CTCTGGGCCT CCCAGTCCTG GGCGCCCCCC CACGCCTCAT CTGTGACAGC 290

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 286 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GTAGCTCGGG GGGTGGGGTG TGCACGCGGC CGCGGGATTG AATGAAGGCA AAGGAGGCAG 60

AGCCTGAGCG CTCGCAAGGT TCGGGTCGGG AGGACTCGCG GGGCCAGAGC AGCGGGATGG 120

GTGAACGTGC AGGTCCAAAC CACCCCTTCT CCTGCAAGTC CCGGCCTCAC ACTCAGCCCA 180

GCTCTCTCTC CTAGAATGTC TTCTGCTGCT TCTGCTGCCC TCCTTGCTGC TGCTTCCTCT 240

GGGCCTCCCA GTCCTGGGCG CCCCCACACA CCTCATCTGT GACAGC 286

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 288 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CCGTAGCTGT TGGGGTGGGG TGCGCACGAA GCCGCGGGAT TGAATGAACG CGGAGGCAGA 60

GCTGAGCGCT CCCAAGGTCC GGGTGGGGAG GGCTCACTGG GGGCTGCGCA GCGGGATGGG 120

TGAACCTGCC GGTCCAAAGC AGCCCTTCTC CTGCGAGAAT GCACCTCATC CGCAGCCCGT 180

CTCTCTTTCC TAGACTGTAC TCCGCTGCTG CTGCTGCTGC TGTCTCTTCT GCTGTTTCCT 240

CTGGGCCTCC CAGTCTTGGG CGCCCCCCCA CGCCTCATCT GTGACAGC 288

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 280 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TGGGGGTGGG GTGTGCAGGC GGCAGCAGGA TTGAATGAAG GCAGAGGAGA CAGACCCTGA 60

GCGCTGGGAA GGTTGGGGGC AGGAGCCACT AGCTGGGGGC AGAGGAGGGG GATGCGTGAA 120

CCTGCCCCTC CAAACCAACC CTCCCCTTCC CCTCCCTGCC TCACACTCAG CCCGGATCTC 180

CTTTCCAGAA TGTCCTGCCC TGCTGCTTCT GCTGTCCCTG CTACTGCCTC CTCTGGGCCT 240

CCCAGCCCTG GGCGCCCCTC CACGCCTCAT CTGTGACACC 280

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 274 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CAGGTCAGTT TGATTGGGTG GATGTGTGCC GCGAAGCAGC GGCCGGATTG AATGAAGGCA 60

GGGGAGGCAG AACCTGAACG CTGGGAAGGT GGGGGTCGGG CGCGACTAGT TGGGGGCAGA 120

GGAGCGGGAT GTGTGAACCT GCCCCTCCAA ACCCACACAG TCAGCCTGGC ACTCTTTTCC 180

AGAATGTCCT GCCCTGTTCC TTTTGCTGTC TTTGCTGCTG CTTCCTCTGG ACCTCCCAGT 240

CCTGGGCGCC CCCCCTCGCC TCATTTGTGA CAGC 274

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 269 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TGGTAGTTTG GCGGGTGGTG TGTGCTCACC GCGGCGGCGG CGGGGATTGA ATGAAGGCCA 60

AGGAGGCAGA ACCCGAGCGC TGGGAAGGTT CGGGGTGGGA GCGACAAGCT GGGGGCAGAG 120

GAACGGGATG TGTGAACCTG CTCCTCCAAA CACACTCAGC CTGGCACTCT TTTCTAGAAT 180

GTCCTGCCCT GCTGCTTCTG CTATCTTTGC TGCTGCTTCC CCTGGGCCTC CCAGTCCTGG 240

GCGCCCCCCC TCGCCTCATC TGTGACAGC 269

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 268 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GGTAGTTTGG CGGGTGGGGT GTGCTCACCG CGGCGGCGGC GGGGATTGAA TGAAGGCCAA 60

GGAGGCAGAA CCCGAGCGCT GGGAAGGTTC GGGGTGGGAG CGACAAGCTG GGGGCAGAGG 120

AACGGGATGT GTGAACCTGC TGCTCCAAAC ACACTCAGCC TGGCACTCTT TTCTAGAATG 180

TCCTGCCCTG CTGCTTCTGC TATCTTTGCT GCTGCTTCCC CTGGGCCTCC CAGTCCTGGG 240

CGCCCCCCCT CGCCTCATCT GTGACAGC 268

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 283 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGGTCGCTTG GGGGGTGGGG TGTGCAGCGC GGAGGGATTG AATGAAGGCC ACTCAGGCAG 60

AGCCTAAGCA ATTGCAAGGT CGGGGTCAGC AGAGACTATA AGGGCAGAGG GGTCTCGCTG 120

AGCCAACCGG CCCCTGCAGT CCCACGGCCC CCTCCCCTCT CGGCCTCACA CTCAGCCTGC 180

CTTTCTTCCA GAACGTCCCA CCCTGCTGCT TTTACTCTCC TTGCTACTGA TTCCTCTGGG 240

CCTCCCAGTC CTCTGTGCTC CCCCACGCCT CATCTGCGAC AGT 283

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 283 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ATGGTGGCTT GGGGGGTGGT GTGTGCAGCG CCGGGAGATA GAATGAAGGC CACTCAGGCA 60

GAGCCTAAGC AGTTGCAAGG TCGGGGTCAG CAGAGACTGG AAGGGCAGAG GAGCCTCGCT 120

GTGCAAACCG GTCTCTGGTC GTCCCACGGC CCTCCCCTCC CAGCCTCACA CCCAGCCTGC 180

CTTTCTTCTA GAACGTCCAA CCCTGCTGCT TTTGCTGTCT TTGCTCCTGC TTCCCCTGGG 240

CCTCCCAGTC CTCTGCGCTC CCCCACGCCT CATCTGCGAC AGC 283

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 112 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GGGTGGTGGC GATGAATGAA GCAGCAGCTA GGTGGGGGGC GGTGAAGTAC AGCCGTCTCA 60

GAGTGTCTTC TCTCTGCCCT GGCCTCCCAG CTGGCCCCCC CTCATTGGAC AG 112

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 725 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

AGGCGCGGAG ATGGGGGTGC ACGAATGTCC TGCCTGGCTG TGGCTTCTCC TGTCCCTGCT 60

GTCGCTCCCT CTGGGCCTCC CAGTCCTGGG CGCCCCACCA CGCCTCATCT GTGACAGCCG 120

AGTCCTGGAG AGGTACCTCT TGGAGGCCAA GGAGGCCGAG AATATCACGA CGGGCTGTGC 180

TGAACACTGC AGCTTGAATG AGAATATCAC TGTCCCAGAC ACCAAAGTTA ATTTCTATGC 240

CTGGAAGAGG ATGGAGGTCG GGCAGCAGGC CGTAGAAGTC TGGCAGGGCC TGGCCCTGCT 300

GTCGGAAGCT GTCCTGCGGG GCCAGGCCCT GTTGGTCAAC TCTTCCCAGC CGTGGGAGCC 360

CCTGCAGCTG CATGTGGATA AAGCCGTCAG TGGCCTTCGC AGCCTCACCA CTCTGCTTCG 420

GGCTCTGGGA GCCCAGAAAG AAGCCATCTC CCCTCCAGAT GCGGCCTCAG CTGCTCCACT 480

CCGAACAATC ACTGCTGACA CTTTCCGCAA ACTCTTCCGA GTCTACTCCA ATTTCCTCCG 540

GGGAAAGCTG AAGCTGTACA CAGGGGAGGC CTGCAGGACA GGGGACAGAT GACCAGGTGT 600

GTCCACCTGG GCATATCCAC CACCTCCCTC ACCAACATTG CTTGTGCCAC ACCCTCCCCC 660

GGCCACTCCT GAACCCCGTC GAGGGGCTCT CAGCTCAGCG CCAGCCTGTC CCATGGACAC 720

TCCAG 725

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 681 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ACGAATGTCC TGCCTGGCTG TGGCTTCTCC TGTCTCTCGT GTCGCTCCCT CTGGGCCTCC 60

CAGTCCCGGG CGCCCCACCA CGCCTCGTCT GTGACAGCCG AGTCCTGGAG AGGTACCTCT 120

TGGAGGCCAA GGAGGCCGAG AATGTCACGA TGGGCTGTTC CGAAAGCTGC AGCTTGAATG 180

AGAATATCAC CGTCCCAGAC ACCAAAGTTA ACTTCTATGC CTGGAAGAGG ATAGAGGTCG 240

GGCAGCAGGC TGTAGAAGTC TGGCAGGGCC TGGCCCTGCT CTCAGAAGCT GTCCTGCGGG 300

GCCAGGCCGT GTTGGCCAAC TCTTCCCAGC CTTTCGAGCC CCTGCAGCTG CACATGGATA 360

AAGCCATCAG TGGCCTTCGC AGCATCACCA CTCTGCTTCG GGCGCTGGGA GCCCAGGAAG 420

CCATCTCCCT CCCAGATGCG GCCTCGGCTG CTCCACTCCG AACCATCACT GCTGACACTT 480

TCTGCAAACT CTTCCGAGTC TACTCCAATT TCCTCCGGGG AAAGCTGAAG CTGTACACGG 540

GGGAGGCCTG CAGGAGAGGG GACAGATGAC CAGGTGCGTC CAGCTGGGCA CATCCACCAA 600

CTCCCTCACC AACACTGCCT GTGCCACACC CTCCCTCACC ACTCCCGAAC CCCATCGAGG 660

GGCTCTCAGC TAAGCGCCAG C 681

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 679 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CCGAACGTCC CACCCTGCTG CTTTTACTCT CCTTGCTACT GATTCCTCTG GGCCTCCCAG 60

TCCTCTGTGC TCCCCCACGC CTCATCTGCG ACAGTCGAGT TCTGGAGAGG TACATCTTAG 120

AGGCCAAGGA GGCAGAAAAT GTCACGATGG GTTGTGCAGA AGGTCCCAGA CTGAGTGAAA 180

ATATTACAGT CCCAGATACC AAAGTCAACT TCTATGCTTG GAAAAGAATG GAGGTGGAAG 240

AACAGGCCAT AGAAGTTTGG CAAGGCCTGT CCCTGCTCTC AGAAGCCATC CTGCAGGCCC 300

AGGCCCTGCT AGCCAATTCC TCCCAGCCAC CAGAGACCCT TCAGCTTCAT ATAGACAAAG 360

CCATCAGTGG TCTACGTAGC CTCACTTCAC TGCTTCGGGT ACTGGGAGCT CAGAAGGAAT 420

TGATGTCGCC TCCAGATACC ACCCCACCTG CTCCACTCCG AACACTCACA GTGGATACTT 480

TCTGCAAGCT CTTCCGGGTC TACGCCAACT TCCTCCGGGG GAAACTGAAG CTGTACACGG 540

GAGAGGTCTG CAGGAGAGGG GACAGGTGAC ATGCTGCTGC CACCGTGGTG GACCGACGAA 600

CTTGCTCCCC GTCACTGTGT CATGCCAACC CTCCACCACT CCCAACCCTC ATCAAACGGG 660

TCATTACCTT CTTACCAGT 679

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 678 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCGAACGTCC CACCCTGCTG CTTTTACTAT CCTTGCTACT GATTCCTCTG GGCCTCCCAG 60

TCCTCTGCGC TCCCCCACGC CTCATTTGCG ACAGTCGCGT TCTGGAGAGG TACATCTTGG 120

AGGCCAAGGA GGCAGAAAAT GTCACAATGG GCTGTGCAGA AGGTCCCAGA CTGAGTGAGA 180

ATATTACCGT CCCAGATACC AAAGTCAACT TCTACGCTTG GAAAAGAATG AAGGTGGAAG 240

AACAGGCTGT AGAAGTTTGG CAAGGCCTGT CTCTGCTCTC AGAAGCCATC CTGCAGGCCC 300

AGGCTCTGCA GGCCAATTCC TCCCAGCCAC CAGAGAGTCT TCAGCTTCAT ATAGACAAAG 360

CCATCAGTGG GCTACGTAGC CTCACTTCAC TGCTTCGGGT GCTGGGAGCT CAGAAGGAAT 420

TGATGTCGCC TCCAGACGCC ACCCAAGCCG CTCCACTCCG AACACTCACA GCGGATACTT 480

TCTGCAAGCT CTTCCGGGTC TACTCCAACT TCCTCCGGGG GAAACTGAAG CTGTACACGG 540

GGGAGGCCTG CAGGAGAGGG GACAGGTGAC CTGCCACTGC CGTGTACCCG CCAACTCGCT 600

CACCGTCACT GTGTCACGCC AACCCTCCAC CACTCCCAAC CCTCATCAAA CGGGGTTGTT 660

TGTTACCTTC TTACCGGC 678

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 687 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCGACTGTAC TCCGCTGCTG CTGCTGCTGC TGTCTCTTCT GCTGTTTCCT CTGGGCCTCC 60

CAGTCTTGGG CGCCCCCCCA CGCCTCATCT GTGACAGCCG AGTCCTGGAG AGGTACATCC 120

TGGAGGCCAG GGAGGCCGAA AATGCCACGA TGGGCTGTGC AGAAGGCTGC AGCTTCAGTG 180

AGAATATCAC TGTCCCAGAC ACCAAGGTTA ACTTCTATGC CTGGAAGAGG ATGGAGGTCC 240

AGCAGCAGGC TCTGGAAGTC TGGCAGGGCC TGGCTCTGCT CTCAGAAGCT ATCTTTCGGG 300

GCCAGGCCCT ACCGGCCAAC GCATCCCAGC CATGCGAGGC CCTGCGGTTG CACGTGGACA 360

AAGCTGTCAG CGGCCTCCGC AGTCTCACCT CCCTGCTTCG GGCGCTGGGA GCCCAGAAAG 420

AAGCCATCCC CCTTCCAGAT GCAACCCCCT CCGCAGCCCC ACTCCGAATA TTCACTGTTG 480

ATGCTTTGTC CAAGCTCTTC CGAATCTACT CCAATTTCCT GAGGGGAAAG CTGACGCTGT 540

ACACAGGGGA GGCCTGCAGG AGAGGGGACA GGTGACCCAG TGCTATCACC CCGGGCACGT 600

CCATCACCTC GCTCACCATC ACTGCCTACG CCATGCCTTC CACGCCGCCA CTCCCAACCC 660

CTGTCGACGA CGGGCCACCA GCTCAGT 687

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 688 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GAATGTCCTG CCCGGCTGCT TCTGCTATCT TTGCTGCTGC TTCCTCTGGG CCTCCCAGTC 60

CTGGGCGCCC CCCCACGCCT CATCTGTGAC AGCCGAGTCC TGGAGAGGTA CATCTTGGAG 120

GCCAAGGAGG GCGAAAATGC CACGATGGGC TGTGCCGAAA GCTGCAGCTT CAGTGAGAAT 180

ATCACTGTCC CAGACACCAA GGTTAACTTC TATGCCTGGA AGAGGATGGA GGTCCAGCAG 240

CAGGCCATGG AGGTCTGGCA GGGTCTGGCC CTGCTCTCAG AAGCCATCCT GCAGGGCCAG 300

GCCCTGTTGG CCAACTCCTC CCAGCCATCT GAGGCCCTGC AGCTGCATGT GGACAAAGCT 360

GTCAGCGGCC TGCGCAGCCT CACCTCCCTG CTTCGGGCGC TGGGAGCCCA GAAGGAAGCC 420

ATCCCCCTTC CAGACGCATC CCCCTCCTCT GCCACCCCAC TCCGAACATT TGCTGTTGAT 480

ACTTTGTGCA AACTTTTCCG CAACTACTCC AATTTCCTGC GGGGAAAGCT GACGCTGTAC 540

ACTGGAGAGG CCTGCAGGAG AAGGGACAGG TGACTTGGTG CTCCCACCCG GGGCATGTCC 600

ACCACTTCGC TCACCACCAC TGCCTGTGCC ATGCCTTCTG CACCTCCACT CCCAACCCCC 660

GCTGAGGGGC CATCAGCTCA GCGCCAGC 688

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 681 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GAATGTCCTG CCCTGCTGCT TCTGCTATCT TTGCTGCTGC TTCCCCTGGG CCTCCCAGTC 60

CTGGGCGCCC CCCCTCGCCT CATCTGTGAC AGCCGAGTCC TGGAGAGGTA CATTCTGGAG 120

GCCAGGGAGG CCGAAAATGT CACGATGGGC TGTGCTGAAG GCTGCAGCTT CAGTGAGAAT 180

ATCACTGTCC CAGACACCAA GGTCAACTTC TATACCTGGA AGAGGATGCA CGTCGGGCAG 240

CAGGCTGTGG AAGTCTGGCA GGGCCTCGCC CTGCTCTCAG AAGCCATCCT GCGGGGCCAG 300

GCCCTGCTGG CCAACTCCTC CCAGCCATCT GAGACCCTGC AGCTGCATGT GGATAAAGCC 360

GTCAGCAGCC TGCGCAGCCT CACCTCCCTG CTTCGGGCAC TGGGAGCCCA GAAGGAAGCC 420

ACCTCCCTTC CAGAGGCAAC CTCTGCTGCT CCACTCCGAA CATTCACTGT CGATACTTTG 480

TGCAAACTTT TCCGAATCTA CTCCAACTTC CTGCGGGGAA AGCTGACGCT GTACACAGGG 540

GAGGCCTGCC GAAGAGGAGA CAGGTGACCA GGTGCTCCTA CCCCGGGCAT GTCCACCACC 600

TCACTTACCA CCAATGCCTG TGCCACGCCC TCTGCACCAC CACTCCTGAC CCCTGTCGGG 660

GGTGATCAGC TCAGCACCAG C 681

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Gly Val His Glu Cys Pro Ala Trp Leu Trp Leu Leu Leu Ser Leu
1 5 10 15

Leu Ser Leu Pro Leu Gly Leu Pro Val Leu Gly Ala Pro Pro
20 25 30

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Met Gly Val His Glu Cys Pro Ala Trp Leu Trp Leu Leu Leu Ser Leu
1 5 10 15

Val Ser Leu Pro Leu Gly Leu Pro Val Pro Gly Ala Pro Pro
20 25 30

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Met Gly Val Pro Glu Arg Pro Thr Leu Leu Leu Leu Leu Ser Leu Leu
1 5 10 15

Leu Ile Pro Leu Gly Leu Pro Val Leu Cys Ala Pro Pro
20 25

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Met Gly Val Arg Asp Cys Thr Pro Leu Leu Leu Leu Leu Leu Ser Leu
1 5 10 15

Leu Leu Phe Pro Leu Gly Leu Pro Val Leu Gly Ala Pro Pro
20 25 30

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Glu Cys Pro Ala Arg Leu Leu Leu Leu Ser Leu Leu Leu Leu Pro Leu

1 5 10 15

Gly Leu Pro Val Leu Gly Ala Pro Pro
20 25

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Glu Cys Pro Ala Leu Leu Leu Leu Leu Ser Leu Leu Leu Pro Pro Leu
1 5 10 15

Gly Leu Pro Ala Leu Gly Ala Pro Pro
20 25

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Glu Cys Pro Ala Leu Leu Leu Leu Leu Ser Leu Leu Leu Leu Pro Leu
1 5 10 15

Gly Leu Pro Val Leu Gly Ala Pro Pro
20 25

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Glu Cys Pro Ala Leu Phe Leu Leu Leu Ser Leu Leu Leu Leu Pro Leu
1 5 10 15

Asp Leu Pro Val Leu Gly Ala Pro Pro
20 25

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```
Glu Cys Leu Leu Leu Leu Leu Leu Pro Ser Leu Leu Leu Leu Pro Leu
1               5                   10                  15

Gly Leu Pro Val Leu Gly Ala Pro Pro
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 166 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

```
Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
1               5                   10                  15

Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His
            20                  25                  30

Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
        35                  40                  45

Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp
    50                  55                  60

Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
65                  70                  75                  80

Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp
                85                  90                  95

Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu
            100                 105                 110

Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
        115                 120                 125

Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val
    130                 135                 140

Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
145                 150                 155                 160

Cys Arg Thr Gly Asp Arg
                165
```

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 165 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

```
Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
1               5                   10                  15

Leu Glu Ala Lys Glu Ala Glu Asn Val Thr Met Gly Cys Ser Glu Ser
            20                  25                  30

Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
        35                  40                  45
```

Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp
50 55 60

Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Val
65 70 75 80

Leu Ala Asn Ser Ser Gln Pro Phe Glu Pro Leu Gln Leu His Met Asp
85 90 95

Lys Ala Ile Ser Gly Leu Arg Ser Ile Thr Thr Leu Leu Arg Ala Leu
100 105 110

Gly Ala Gln Glu Ala Ile Ser Leu Pro Asp Ala Ala Ser Ala Ala Pro
115 120 125

Leu Arg Thr Ile Thr Ala Asp Thr Phe Cys Lys Leu Phe Arg Val Tyr
130 135 140

Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys
145 150 155 160

Arg Arg Gly Asp Arg
165

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 165 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Ala Pro Pro Arg Leu Val Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
1 5 10 15

Leu Glu Ala Lys Glu Ala Glu Asn Val Thr Met Gly Cys Ser Glu Ser
20 25 30

Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
35 40 45

Tyr Ala Trp Lys Arg Ile Glu Val Gly Gln Gln Ala Val Glu Val Trp
50 55 60

Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Val
65 70 75 80

Leu Ala Asn Ser Ser Gln Pro Phe Glu Pro Leu Gln Leu His Met Asp
85 90 95

Lys Ala Ile Ser Gly Leu Arg Ser Ile Thr Thr Leu Leu Arg Ala Leu
100 105 110

Gly Ala Gln Glu Ala Ile Ser Leu Pro Asp Ala Ala Ser Ala Ala Pro
115 120 125

Leu Arg Thr Ile Thr Ala Asp Thr Phe Cys Lys Leu Phe Arg Val Tyr
130 135 140

Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys
145 150 155 160

Arg Arg Gly Asp Arg
165

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 166 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Ile
1 5 10 15

Leu Glu Ala Lys Glu Ala Glu Asn Val Thr Met Gly Cys Ala Glu Gly
20 25 30

Pro Arg Leu Ser Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
35 40 45

Tyr Ala Trp Lys Arg Met Glu Val Glu Glu Gln Ala Ile Glu Val Trp
50 55 60

Gln Gly Leu Ser Leu Leu Ser Glu Ala Ile Leu Gln Ala Gln Ala Leu
65 70 75 80

Leu Ala Asn Ser Ser Gln Pro Pro Glu Thr Leu Gln Leu His Ile Asp
85 90 95

Lys Ala Ile Ser Gly Leu Arg Ser Leu Thr Ser Leu Leu Arg Val Leu
100 105 110

Gly Ala Gln Lys Glu Leu Met Ser Pro Pro Asp Thr Thr Pro Pro Ala
115 120 125

Pro Leu Arg Thr Leu Thr Val Asp Thr Phe Cys Lys Leu Phe Arg Val
130 135 140

Tyr Ala Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Val
145 150 155 160

Cys Arg Arg Gly Asp Arg
165

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 166 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Ile
1 5 10 15

Leu Glu Ala Lys Glu Ala Glu Asn Val Thr Met Gly Cys Ala Glu Gly
20 25 30

Pro Arg Leu Ser Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
35 40 45

Tyr Ala Trp Lys Arg Met Lys Val Glu Glu Gln Ala Val Glu Val Trp
50 55 60

Gln Gly Leu Ser Leu Leu Ser Glu Ala Ile Leu Gln Ala Gln Ala Leu

65 70 75 80

Gln Ala Asn Ser Ser Gln Pro Pro Glu Ser Leu Gln Leu His Ile Asp
85 90 95

Lys Ala Ile Ser Gly Leu Arg Ser Leu Thr Ser Leu Leu Arg Val Leu
100 105 110

Gly Ala Gln Lys Glu Leu Met Ser Pro Pro Asp Ala Thr Gln Ala Ala
115 120 125

Pro Leu Arg Thr Leu Thr Ala Asp Thr Phe Cys Lys Leu Phe Arg Val
130 135 140

Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
145 150 155 160

Cys Arg Arg Gly Asp Arg
165

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 167 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Ile
1 5 10 15

Leu Glu Ala Arg Glu Ala Glu Asn Ala Thr Met Gly Cys Ala Glu Gly
20 25 30

Cys Ser Phe Ser Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
35 40 45

Tyr Ala Trp Lys Arg Met Glu Val Gln Gln Gln Ala Leu Glu Val Trp
50 55 60

Gln Gly Leu Ala Leu Leu Ser Glu Ala Ile Phe Arg Gly Gln Ala Leu
65 70 75 80

Pro Ala Asn Ala Ser Gln Pro Cys Glu Ala Leu Arg Leu His Val Asp
85 90 95

Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Ser Leu Leu Arg Ala Leu
100 105 110

Gly Ala Gln Lys Glu Ala Ile Pro Leu Pro Asp Ala Thr Pro Ser Ala
115 120 125

Ala Pro Leu Arg Ile Phe Thr Val Asp Ala Leu Ser Lys Leu Phe Arg
130 135 140

Ile Tyr Ser Asn Phe Leu Arg Gly Lys Leu Thr Leu Tyr Thr Gly Glu
145 150 155 160

Ala Cys Arg Arg Gly Asp Arg
165

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 168 amino acids
(B) TYPE: amino acid

(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Ile
1 5 10 15

Leu Glu Ala Lys Glu Gly Glu Asn Ala Thr Met Gly Cys Ala Glu Ser
20 25 30

Cys Ser Phe Ser Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
35 40 45

Tyr Ala Trp Lys Arg Met Glu Val Gln Gln Gln Ala Met Glu Val Trp
50 55 60

Gln Gly Leu Ala Leu Leu Ser Glu Ala Ile Leu Gln Gly Gln Ala Leu
65 70 75 80

Leu Ala Asn Ser Ser Gln Pro Ser Glu Ala Leu Gln Leu His Val Asp
85 90 95

Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Ser Leu Leu Arg Ala Leu
100 105 110

Gly Ala Gln Lys Glu Ala Ile Pro Leu Pro Asp Ala Ser Pro Ser Ser
115 120 125

Ala Thr Pro Leu Arg Thr Phe Ala Val Asp Thr Leu Cys Lys Leu Phe
130 135 140

Arg Asn Tyr Ser Asn Phe Leu Arg Gly Lys Leu Thr Leu Tyr Thr Gly
145 150 155 160

Glu Ala Cys Arg Arg Arg Asp Arg
165

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 166 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Ile
1 5 10 15

Leu Glu Ala Arg Glu Ala Glu Asn Val Thr Met Gly Cys Ala Glu Gly
20 25 30

Cys Ser Phe Ser Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
35 40 45

Tyr Thr Trp Lys Arg Met Asp Val Gly Gln Gln Ala Val Glu Val Trp
50 55 60

Gln Gly Leu Ala Leu Leu Ser Glu Ala Ile Leu Arg Gly Gln Ala Leu
65 70 75 80

Leu Ala Asn Ser Ser Gln Pro Ser Glu Thr Leu Gln Leu His Val Asp
85 90 95

Lys Ala Val Ser Ser Leu Arg Ser Leu Thr Ser Leu Leu Arg Ala Leu
100 105 110

Gly Ala Gln Lys Glu Ala Thr Ser Leu Pro Glu Ala Thr Ser Ala Ala
115 120 125

Pro Leu Arg Thr Phe Thr Val Asp Thr Leu Cys Lys Leu Phe Arg Ile
130 135 140

Tyr Ser Asn Phe Leu Arg Gly Lys Leu Thr Leu Tyr Thr Gly Glu Ala
145 150 155 160

Cys Arg Arg Gly Asp Arg
165

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 166 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu
1 5 10 15

Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His
20 25 30

Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe
35 40 45

Tyr Ala Trp Lys Arg Met Glu Gly Val Gln Gln Ala Val Glu Val Trp
50 55 60

Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu
65 70 75 80

Leu Val Asn Ser Ser Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp
85 90 95

Lys Ala Val Ser Gly Leu Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu
100 105 110

Gly Ala Gln Lys Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala
115 120 125

Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val
130 135 140

Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala
145 150 155 160

Cys Arg Thr Gly Asp Arg
165

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Xaa Xaa Ala Lys
1 5 10

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val
1 5 10 15

Leu Arg

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu
1 5 10 15

Thr Thr

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn
1 5 10 15

Phe Leu Arg Gly Lys
20

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn
1 5 10 15

Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Asn
20 25 30

Glu Val Gly Gln
35

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser
1 5 10 15

Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg
20 25

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Leu Ile Ser Glu
1 5 10 15

Glu Asp Leu Glu Ala Ile Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro
20 25 30

Leu Arg

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Cys Arg Thr Gly Asp Arg
1 5

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Lys Asp Glu Leu
1

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

His His His His His His
1 5

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Met Gly His Ser Ser Gly His Ile Glu Gly Arg His Met Ala Pro Pro
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu Ala
1 5 10

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Arg Thr Ile Thr Ala Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser
1 5 10 15

Asn Phe Leu Arg
20

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: both

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys
1 5 10

Claims:

1. An erythropoietin mutein comprising native erythropoietin with one or more of the following modifications:

(a) replacement of the amino acid at position 20 with a first substitute amino acid;

(b) replacement of the amino acid at position 49 with a second substitute amino acid;

(c) replacement of the amino acid at position 73 with a third substitute amino acid;

(d) replacement of the amino acid at position 140 with a fourth substitute amino acid;

(e) replacement of the amino acid at position 143 with a fifth substitute amino acid; and

(f) replacement of the amino acid at position 146 with a sixth substitute amino acid;

(g) replacement of the amino acid at position 147 with a seventh substitute amino acid;

(h) replacement of the amino acid at position 154 with a eighth substitute amino acid;

wherein said erythropoietin mutein exhibits enhanced biological activity compared to said native erythropoietin.

2. The erythropoietin mutein of claim 1 comprising native erythropoietin with one of said modifications.

3. The erythropoietin mutein of claim 1 wherein said first, second, third, fourth, fifth, sixth, seventh or eighth substitute amino acids are selected from the group consisting of alanine, serine, threonine, and glycine.

4. The erythropoietin mutein of claim 3 wherein said first substitute amino acid is serine, and/or said second, third, fourth, fifth, sixth, seventh and/or eighth substitute amino acids are alanine.

5. The erythropoietin mutein of claim 2 wherein the amino acid at position 20 is replaced with the amino acid alanine.

6. The erythropoietin mutein of claim 2 wherein the amino acid at position 49 is replaced with the amino acid serine.

7. The erythropoietin mutein of claim 2 wherein the amino acid at position 73 is replaced with the amino acid glycine.

8. The erythropoietin mutein of claim 2 wherein the amino acid at position 140 is replaced with the amino acid alanine.

9. The erythropoietin mutein of claim 2 wherein the amino acid at position 143 is replaced with the amino acid alanine.

10. The erythropoietin mutein of claim 2 wherein the amino acid at position 146 is replaced with the amino acid alanine.

11. The erythropoietin mutein of claim 2 wherein the amino acid at position 147 is replaced with the amino acid alanine.

12. The erythropoietin mutein of claim 2 wherein the amino acid at position 154 is replaced with the amino acid alanine.

13. A DNA molecule encoding the erythropoietin mutein of claim 1.

14. A recombinant DNA construct comprising the DNA molecule of claim 13.

15. A vector comprising the DNA construct of claim 14 which is capable of expressing said erythropoietin mutein in a host cell.

16. A host cell comprising the DNA molecule of claim 13 capable of expressing the erythropoietin mutein encoded by said DNA molecule.

17. The host cell of claim 16 wherein said host cell is capable of glycosylating said erythropoietin mutein.

18. The host cell of claim 17 wherein said host cell is a mammalian cell.

19. The host cell of claim 18 wherein said host cell is selected from the group consisting of Chinese Hamster Ovary cells, Cos7 cells,Cos1 cells, baby hamster kidney cells, and CV1 cells.

20. A method of stimulating erythrocyte production in a subject, wherein said method comprises administering to said subject an efficacious amount of the erythropoietin mutein of claim 1.

21. A method of stimulating erythrocyte production in a subject, said method comprising administering to said subject an efficacious amount of an erythropoietin mutein comprising native erythropoietin with the amino acid residue at position 143 replaced by an alanine residue, wherein said efficacious amount is less than the corresponding amount of native erythropoietin.

22. A method of inducing the proliferation of a cell culture responsive to erythropoietin, wherein said method comprises contacting said cell culture with the erythropoietin mutein of claim 1.

23. A method of inducing the proliferation of a cell culture responsive to erythropoietin, wherein said method comprises contacting said cell culture with an efficacious amount of an erythropoietin mutein comprising native erythropoietin with the amino acid residue at position 143 replaced by an alanine residue, wherein said efficacious amount is less than the corresponding amount of native erythropoietin.

24. The method of claim 18 or 19 wherein said cell culture comprises cells selected from the group consisting of murine spleen cells, murine erythroleukemia cells, and human UT-7/Epo cells.

25. A method for obtaining erythropoietin muteins with enhanced biological activity relative to native erythropoietin, said method comprising the steps:

(a) preparing erythropoietin muteins comprising native erythropoietin with a single amino acid substitution, said substitution occurring at a position selected from the group consisting of positions 48-52, 151-160, and positions predicted to reside on the external surface of helices A and D;

(b) assaying the biological activity of the erythropoietin muteins from step (a) over a range of dosages; and

(c) identifying those erythropoietin muteins with enhanced biological activity relative to native erythropoietin.

			localization (bp) human (1)	mouse(2)
IV1	5'	TGAAGTTTGGCCGAGAAGTGGATGC 3'	984–1009	994–1018
EX2R	5'	AAGA(T/G)GTACCTCTCCAG(A/G)ACTCG 3'	1296–1318	1302–1324
EX5	5'	CTGCTCCACTCCGAACA(C/A)TCAC 3'	2651–2672	3183–3204
NCOI	5'	CTGGAGTGTCCATGGGACAG 3'	2885–2904	3415–3434
ATG	5'	AGGCGCGGAGATGGGGGTGC 3'	615–634	633–652

FIG.1

	-27																				1		
Human	M	G	V	H	E	C	P	A	W	L	W	L	L	S	L	L	S	L	P	L	G	L	P	V	L	G↑A	P	P
Monkey	M	G	V	H	E	C	P	A	W	L	W	L	L	S	L	V	S	L	P	L	G	L	P↑V	P	G	A	P	P
Rodents	M	G	V	P	E	R	P	T	-	L	L	L	L	S	L	L	L	I	P	L	G	L	P	V	L	C	A	P	P
Sheep	M	G	V	R	D	C	T	P	L	L	L	L	L	S	L	L	L	F	P	L	G	L	P	V	L	G	A	P	P
Pig					E	C	P	A	R	-	L	L	L	S	L	L	L	L	P	L	G	L	P	V	L	G	A	P	P
Horse					E	C	P	A	-	L	L	L	L	S	L	L	L	P	P	L	G	L	P	A	L	G	A	P	P
Cat					E	C	P	A	-	L	L	L	L	S	L	L	L	L	P	L	G	L	P	V	L	G	A	P	P
Dog					E	C	P	A	-	L	F	L	L	S	L	L	L	L	P	L	D	L	P	V	L	G	A	P	P
Dolphin					E	C	-	L	L	L	L	L	L	P	S	L	L	L	L	P	L	G	L	P	V	L	G	A	P	P

FIG.5

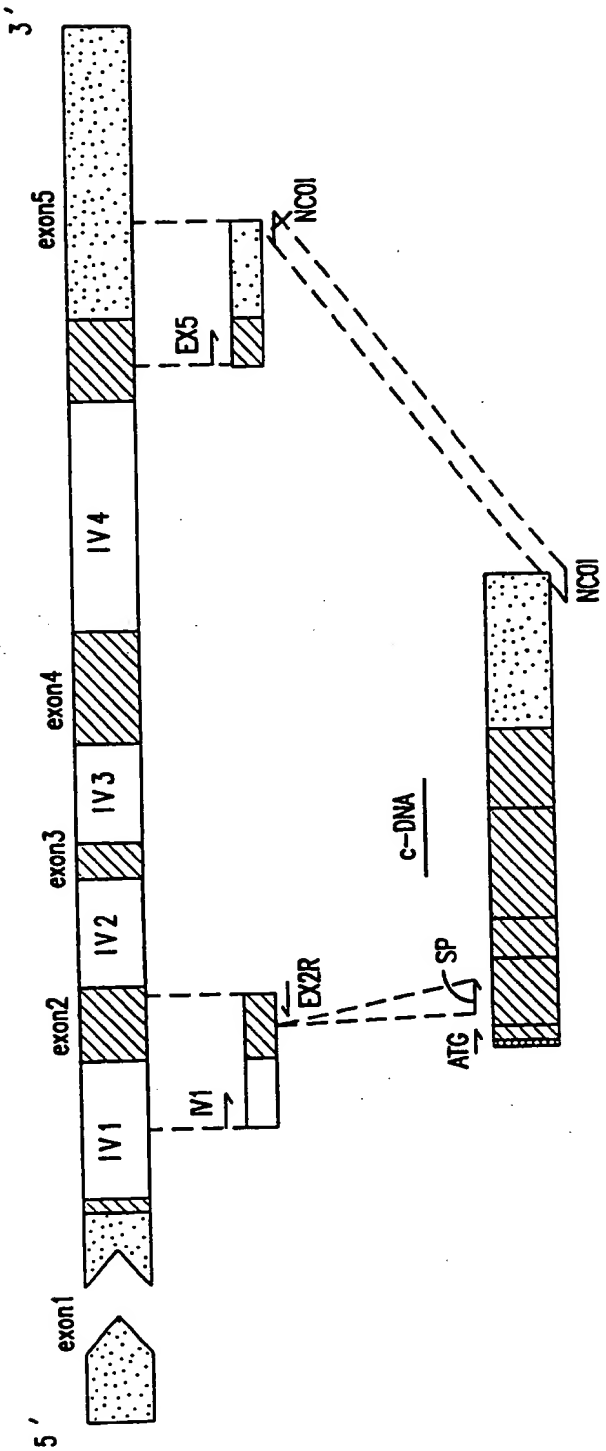

3'
exon5
NCOI
EX5
IV4
exon4
IV3
exon3
IV2
exon2
EX2R
SP
c-DNA
ATG
IV1
exon1
5'
NCOI

FIG.2

```
               1010      1020       1030          1040       1050      1060      1070
         IV1
Human ------>  TGGTAGCT   GGGGGTGGGGTGTG  CAC ACGG     CAGCAGGATTGAATGAAGGCCAGGGAGGCAGCAC
Pig           CTCGTAGTT GTT GGGTGGCGTGTG  CAC GCGG     TCAGTGTATTGAATGAAGGCAAAGGAGGCAGAGC
Dolphin          GTAGCTC GGGGGGTGGGGTGTG  CAC GCGG     CCGCGGGATTGAATGAAGGCAAAGGAGGCAGAGC
Sheep          CCGTAGCT GTTGCGGTGGGGTGCG  CAC GAAG     CCGCGGGATTGAATGAACGC   GGAGGCAGAGC
Horse                   TGGGGGTGGGGTGTG  CAG GCGGCAG     CAGGATTGAATGAAGGCAGAGGAGACAGACC
Dog           CAGGTCAGTTTGATTGGGTGG ATGTGTGCCGCGAAGCAGCGG CCGGATTGAATGAAGGCAGGGGAGGCAGAAC
Cat            TGGTAGTTT GGCGGGTGGTGTGTGCTCACCGCGGCGGCGGCGGGGATTGAATGAAGGCCAAGGAGGCAGAAC
Lion            GGTAGTTT GGCGGGTGGGGTGTGCTCACCGCGGCGGCGGCGGGGATTGAATGAAGGCCAAGGAGGCAGAAC
Mouse          CGGTCGCTT GGGGGGTGGGGTGTG  CAGCGCGG        AGGGATTGAATGAAGGCCACTCAGGCAGAGC
Hamster       ATGGTGGCTT GGGGGGTGGTGTGTG  CAGCGCCG        GGAGATAGAATGAAGGCCACTCAGGCAGAGC
Consensus     ---------------GGGTGG--TG-G--C------G-------------AT-GAATGAA-GC------AG-CAG--C

               1080      1090      1100      1110       1120       1130      1140
Human         CTGAGTGCTTGCATGGTTGGGGACAGGAAGGACGAGCTGGGG CAGAGACG  TGGGATGAAGGAAGCTGTCCTTCCAC
Pig           CTGCGCGCTCGCAAGGTTGGGGTCGGAAACGGCTGGCTGGGGGCAA GACGGCCGGATGGGTGAACCTGTCCGTCTAA
Dolphin       CTGAGCGCTCGCAAGGTTCGGGTCGGGAGGAC TCGC GGGGCCAGAGCAG  CGGGATGGGTGAACGTGCAGGTCCAA
Sheep          TGAGCGCTCCCAAGGTCCGGGTGGGGAGGGC TCACTGGGGGCTGCGCAG  CGGGATGGGTGAACCTGCCGGTCCAA
Horse         CTGAGCGCTGGGAAGGTTGGGGGCAGGAGCCACTAGCTGGGGGCAGAGGAG  GGGGATGCGTGAACCTGCCCCTCCAA
Dog           CTGAACGCTGGGAAGGTGGGGGTCGGGCGCGACTAGTTGGGGGCAGAGGAG  CGGGATGTGTGAACCTGCCCCTCCAA
Cat           CCGAGCGCTGGGAAGGTTCGGGGTGGGAGCGACAAGCTGGGGGCAGAGGAA  CGGGATGTGTGAACCTGCTCCTCCAA
Lion          CCGAGCGCTGGGAAGGTTCGGGGTGGGAGCGACAAGCTGGGGGCAGAGGAA  CGGGATGTGTGAACCTGCTGCTCCAA
Mouse         CTAAGCAATTGCAAGGTCGGGGTCAGCAGAGACTAT   AAGGGCAGAGGGGTCTCGCTGAGCCAACCGGCCCCTGCAG
Hamster       CTAAGCAGTTGCAAGGTCGGGGTCAGCAGAGACTGG   AAGGGCAGAGGAGCCTCGCTGTGCAAACCGGTCTCTGGTC
Consensus     --------T----A-GGT--GGG---G--------------GG-C---G--  -----G-TG------AA----G-----T------
```

FIG.3

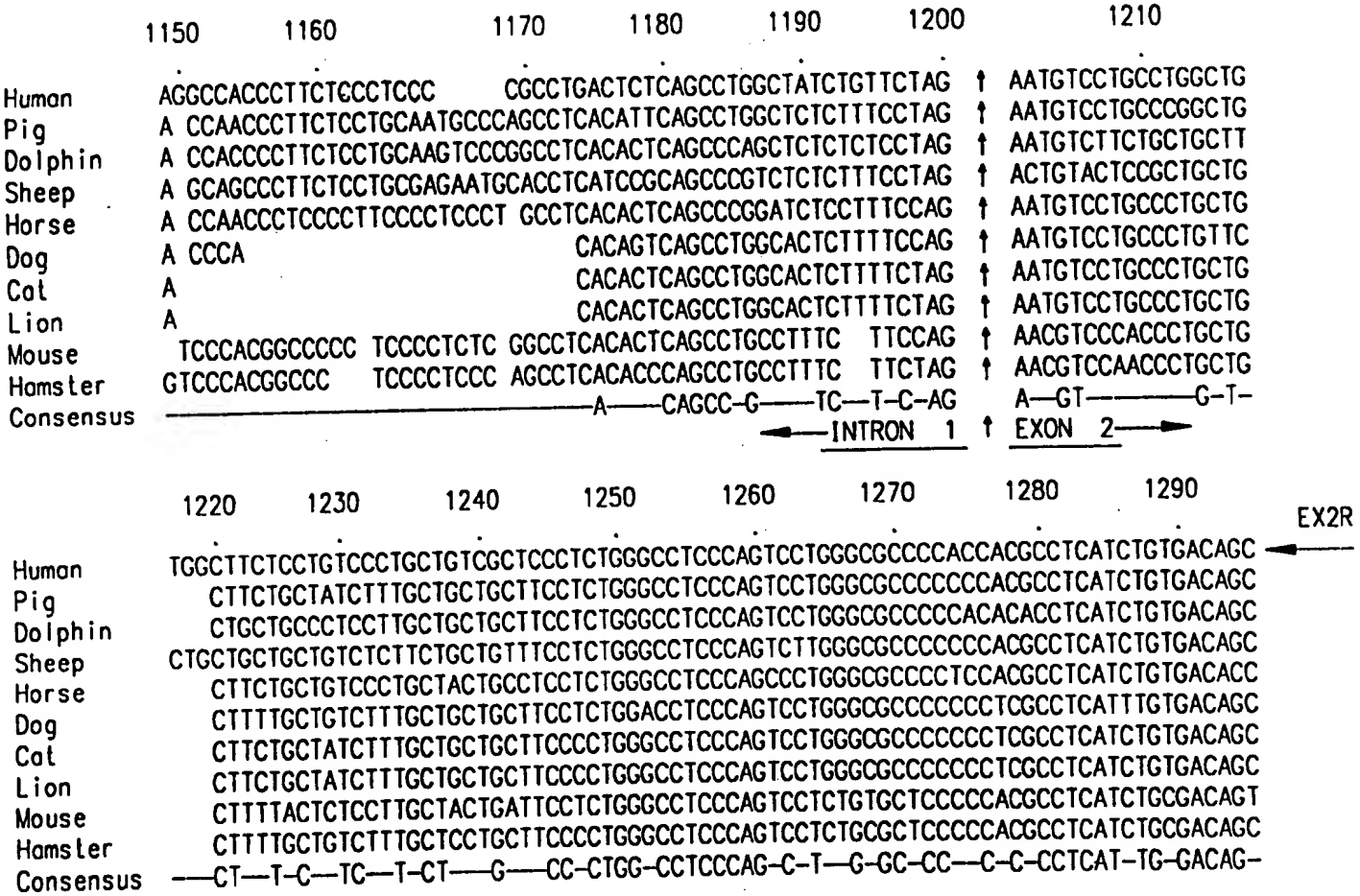


FIG.3A

```
                                                                  50
       1
Human  AGGCGCGGAG ATGGGGTGC ACGAATGTCC TGCCTGGCTG TGGCTTCTCC TGTCCCTGCT GTCGCTCCCT CTGGGCCTCC CAGTCCTGGG
Monkey ---------- --------- ACGAATGTCC TGCCTGGCTG TGGCTTCTCC TGTCTCTCGT GTCGCTCCCT CTGGGCCTCC CAGTCCCGGG
Mouse  ---------- --------- CCGAACGTCC CACC---CTG CTGCTTTTAC TCTCCTTGCT ACTGATTCCT CTGGGCCTCC CAGTCCTCTG
Rat    ---------- --------- CCGAACGTCC CACC---CTG CTGCTTTTAC TATCCTTGCT ACTGATTCCT CTGGGCCTCC CAGTCCTCTG
Sheep  ---------- --------- GCGACTGTAC TCCGCTGCTG CTGCTGCTGC TGTCTCTTCT GCTGTTTCCT CTGGGCCTCC CAGTCTTGGG
Pig                           GAATGTCC TGCCCGG--- CTGCTTCTGC TATCTTTGCT GCTGCTTCCT CTGGGCCTCC CAGTCCTGGG
Cat                           GAATGTCC TGCC---CTG CTGCTTCTGC TATCTTTGCT GCTGCTTCCC CTGGGCCTCC CAGTCCTGGG

    100                                                     150
CGCCCCACCA CGCCTCATCT GTGACAGCCG AGTCCTGGAG AGGTACCTCT TGGAGGCCAA GGAGGCCGAG AATATCACGA CGGGCTGTGC
CGCCCCACCA CGCCTCGTCT GTGACAGCCG AGTCCTGGAG AGGTACCTCT TGGAGGCCAA GGAGGCCGAG AATGTCACGA TGGGCTGTTC
TGCTCCCCCA CGCCTCATCT GCGACAGTCG AGTTCTGGAG AGGTACATCT TAGAGGCCAA GGAGGCAGAA AATGTCACGA TGGGTTGTGC
CGCTCCCCCA CGCCTCATTT GCGACAGTCG CGTTCTGGAG AGGTACATCT TGGAGGCCAA GGAGGCAGAA AATGTCACAA TGGGCTGTGC
CGCCCCCCCA CGCCTCATCT GTGACAGCCG AGTCCTGGAG AGGTACATCC TGGAGGCCAG GGAGGCCGAA AATGCCACGA TGGGCTGTGC
CGCCCCCCCA CGCCTCATCT GTGACAGCCG AGTCCTGGAG AGGTACATCT TGGAGGCCAA GGAGGGCGAA AATGCCACGA TGGGCTGTGC
CGCCCCCCCT CGCCTCATCT GTGACAGCCG AGTCCTGGAG AGGTACATTC TGGAGGCCAG GGAGGCCGAA AATGTCACGA TGGGCTGTGC

                   200                                                   250
TGAACACTGC AGCTTGAATG AGAATATCAC TGTCCCAGAC ACCAAAGTTA ATTTCTATGC CTGGAAGAGG ATGGAGGTCG GGCAGCAGGC
CGAAAGCTGC AGCTTGAATG AGAATATCAC CGTCCCAGAC ACCAAAGTTA ACTTCTATGC CTGGAAGAGG ATAGAGGTCG GGCAGCAGGC
AGAAGGTCCC AGACTGAGTG AAAATATTAC AGTCCCAGAT ACCAAAGTCA ACTTCTATGC TTGGAAAAGA ATGGAGGTGG AAGAACAGGC
AGAAGGTCCC AGACTGAGTG AGAATATTAC CGTCCCAGAT ACCAAAGTCA ACTTCTACGC TTGGAAAAGA ATGAAGGTGG AAGAACAGGC
AGAAGGCTGC AGCTTCAGTG AGAATATCAC TGTCCCAGAC ACCAAGGTTA ACTTCTATGC CTGGAAGAGG ATGGAGGTCC AGCAGCAGGC
CGAAAGCTGC AGCTTCAGTG AGAATATCAC TGTCCCAGAC ACCAAGGTTA ACTTCTATGC CTGGAAGAGG ATGGAGGTCC AGCAGCAGGC
TGAAGGCTGC AGCTTCAGTG AGAATATCAC TGTCCCAGAC ACCAAGGTCA ACTTCTATAC CTGGAAGAGG ATGCACGTCG GGCAGCAGGC
```

FIG.4

```
                                300                                                                                        350
CGTAGAAGTC TGGCAGGGCC TGGCCCTGCT GTCGGAAGCT GTCCTGCGGG GCCAGGCCCT GTTGGTCAAC TCTTCCCAGC CGTGGGAGCC
TGTAGAAGTC TGGCAGGGCC TGGCCCTGCT CTCAGAAGCT GTCCTGCGGG GCCAGGCCGT GTTGGCCAAC TCTTCCCAGC CTTTCGAGCC
CATAGAAGTT TGGCAAGGCC TGTCCCTGCT CTCAGAAGCC ATCCTGCAGG CCCAGGCCCT GCTAGCCAAT TCCTCCCAGC CACCAGAGAC
TGTAGAAGTT TGGCAAGGCC TGTCTCTGCT CTCAGAAGCC ATCCTGCAGG CCCAGGCTCT GCAGGCCAAT TCCTCCCAGC CACCAGAGAG
TCTGGAAGTC TGGCAGGGCC TGGCTCTGCT CTCAGAAGCT ATCTTTCGGG GCCAGGCCCT ACCGGCCAAC GCATCCCAGC CATGCGAGGC
CATGGAGGTC TGGCAGGGTC TGGCCCTGCT CTCAGAAGCC ATCCTGCAGG GCCAGGCCCT GTTGGCCAAC TCCTCCCAGC CATCTGAGGC
TGTGGAAGTC TGGCAGGGCC TCGCCCTGCT CTCAGAAGCC ATCCTGCGGG GCCAGGCCCT GCTGGCCAAC TCCTCCCAGC CATCTGAGAC

                                                 400                                                                        450
CCTGCAGCTG CATGTGGATA AAGCCGTCAG TGGCCTTCGC AGCCTCACCA CTCTGCTTCG GGCTCTGGGA GCCCAGAAAG AAGCCATCTC
CCTGCAGCTG CACATGGATA AAGCCATCAG TGGCCTTCGC AGCATCACCA CTCTGCTTCG GGCGCTGGGA GCCCAG———G AAGCCATCTC
CCTTCAGCTT CATATAGACA AAGCCATCAG TGGTCTACGT AGCCTCACTT CACTGCTTCG GGTACTGGGA GCTCAGAAGG AATTGATGTC
TCTTCAGCTT CATATAGACA AAGCCATCAG TGGGCTACGT AGCCTCACTT CACTGCTTCG GGTGCTGGGA GCTCAGAAGG AATTGATGTC
CCTGCGGTTG CACGTGGACA AAGCTGTCAG CGGCCTCCGC AGTCTCACCT CCCTGCTTCG GGCGCTGGGA GCCCAGAAAG AAGCCATCCC
CCTGCAGCTG CATGTGGACA AAGCTGTCAG CGGCCTGCGC AGCCTCACCT CCCTGCTTCG GGCGCTGGGA GCCCAGAAGG AAGCCATCCC
CCTGCAGCTG CATGTGGATA AAGCCGTCAG CAGCCTGCGC AGCCTCACCT CCCTGCTTCG GGCACTGGGA GCCCAGAAGG AAGCCACCTC

                                                            500
CCCTCCAGAT GCGGCC———T CAGCTG———C TCCACTCCGA ACAATCACTG CTGACACTTT CCGCAAACTC TTCCGAGTCT ACTCCAATTT
CCTCCCAGAT GCGGCC———T CGGCTG———C TCCACTCCGA ACCATCACTG CTGACACTTT CTGCAAACTC TTCCGAGTCT ACTCCAATTT
GCCTCCAGAT ACCACC———C CACCTG———C TCCACTCCGA ACACTCACAG TGGATACTTT CTGCAAGCTC TTCCGGGTCT ACGCCAACTT
GCCTCCAGAC GCCACC———C AAGCCG———C TCCACTCCGA ACACTCACAG CGGATACTTT CTGCAAGCTC TTCCGGGTCT ACTCCAACTT
CCTTCCAGAT GCAACCCCCT CCGCAG———C CCCACTCCGA ATATTCACTG TTGATGCTTT GTCCAAGCTC TTCCGAATCT ACTCCAATTT
CCTTCCAGAC GCATCCCCCT CCTCTGCCAC CCCACTCCGA ACATTTGCTG TTGATACTTT GTGCAAACTT TTCCGCAACT ACTCCAATTT
CCTTCCAGAG GCAACC———T CTGCTG———C TCCACTCCGA ACATTCACTG TCGATACTTT GTGCAAACTT TTCCGAATCT ACTCCAACTT
```

FIG.4A

```
      550                                                                  600
CCTCCGGGGA AAGCTGAAGC TGTACACAGG GGAGGCCTGC AGGACAGGGG ACAGATGACC AGGTGTGTCC ACC-TGGGCA TATCCACCAC
CCTCCGGGGA AAGCTGAAGC TGTACACGGG GGAGGCCTGC AGGAGAGGGG ACAGATGACC AGGTGCGTCC AGC-TGGGCA CATCCACCAA
CCTCCGGGGG AAACTGAAGC TGTACACGGG AGAGGTCTGC AGGAGAGGGG ACAGGTGACA TGCTGCTGCC ACC-GTGGTG GACCGACGAA
CCTCCGGGGG AAACTGAAGC TGTACACGGG GGAGGCCTGC AGGAGAGGGG ACAGGTGACC TGC-------C ACT-GCCGTG TACCCGCCAA
CCTGAGGGGA AAGCTGACGC TGTACACAGG GGAGGCCTGC AGGAGAGGGG ACAGGTGACC CAGTGCTATC ACCCCGGGCA CGTCCATCAC
CCTGCGGGGA AAGCTGACGC TGTACACTGG AGAGGCCTGC AGGAGAAGGG ACAGGTGACT TGGTGCTCCC ACCCGGGGCA TGTCCACCAC
CCTGCGGGGA AAGCTGACGC TGTACACAGG GGAGGCCTGC CGAAGAGGAG ACAGGTGACC AGGTGCTCCT ACCCCGGGCA TGTCCACCAC

                650                                                                  700
CTCCCTCACC AACATTGCTT GTGCCACACC CTCCCCCG-- GCCACTCCTG AACCCC—GT CGA—GGGG- ——CTCTCAG CTCAGCGCCA
CTCCCTCACC AACACTGCCT GTGCCACACC CTCCCTC—— ACCACTCCCG AACCCC—AT CGA——GGGG ——CTCTCAG CTAAGCGCCA
CTTGCTCCCC GTCACTGTGT CATGCCAACC CTCC------ ACCACTCCC- AACCC-TCAT CAAAC—GGG ——TCATTAC CTTCTTACCA
CTCGCTCACC GTCACTGTGT CACGCCAACC CTCC------ ACCACTCCC- AACCC-TCAT CAAACGGGGT TGTTTGTTAC CTTCTTACCG
CTCGCTCACC ATCACTGCCT -ACGCCATGC CTTCCACGCC GCCACTCCC- AACCCC-TGT CGACGACGGG ——CCACCAG CTCAGT
TTCGCTCACC ACCACTGCCT GTGCCATGCC TTCT-GCACC TCCACTCCC- AACCCCC-GC TGA——GGGG ——CCATCAG CTCAGCGCCA
CTCACTTACC ACCAATGCCT GTGCCACGCC CTCT-GCACC ACCACTCCTG A-CCCC-TGT C——GGGGG TG——ATCAG CTCAGCACCA

                    742
GCCTGTCCCA TGGACACTCC AG
GC--------- ---------- --
GT--------- ---------- --
GC--------- ---------- --
  --------- ---------- --
GC--------- ---------- --
GC--------- ---------- --
```

FIG.4B

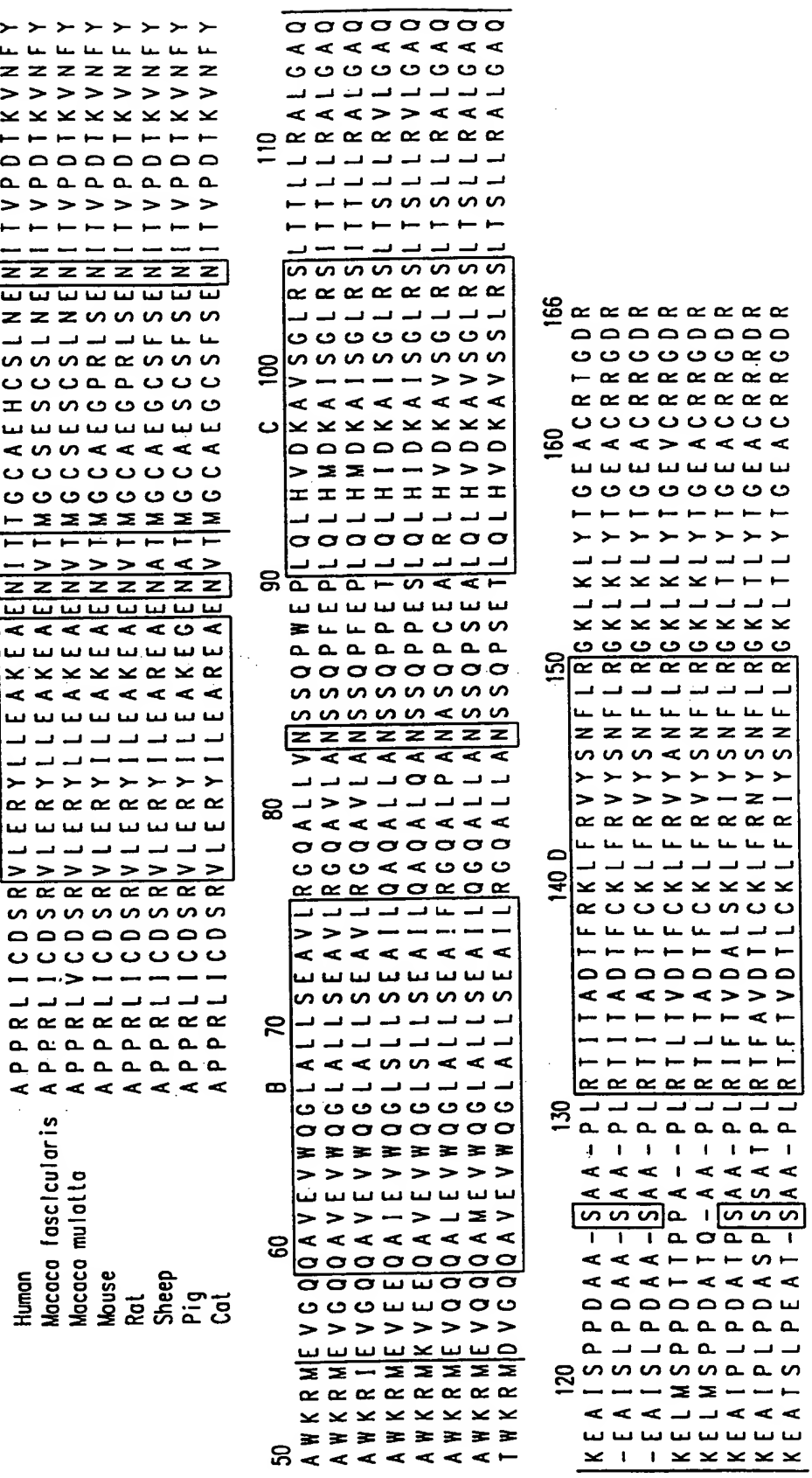

Human
Macaca fascicularis
Macaca mulatta
Mouse
Rat
Sheep
Pig
Cat
A
B
C
D

FIG.6

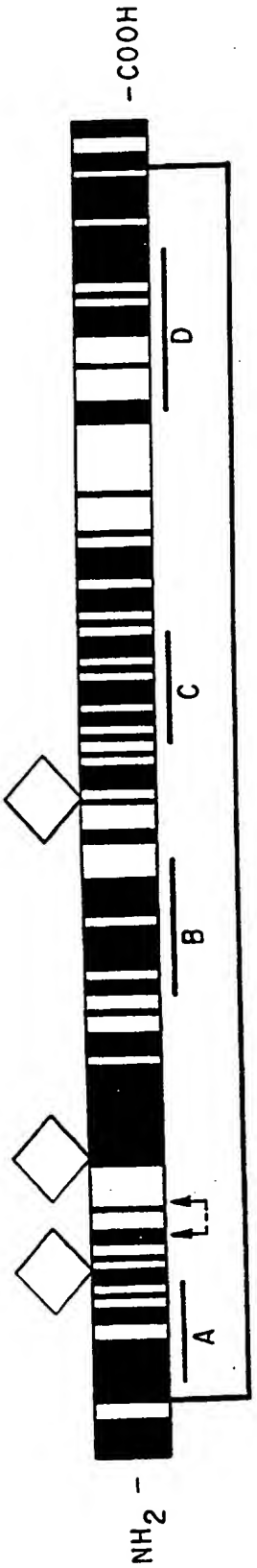


FIG. 6A

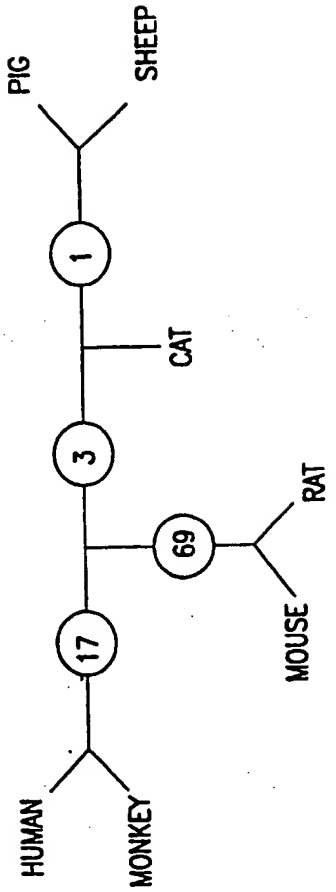

PIG
SHEEP
1
CAT
3
69
RAT
MOUSE
17
HUMAN
MONKEY

FIG.7

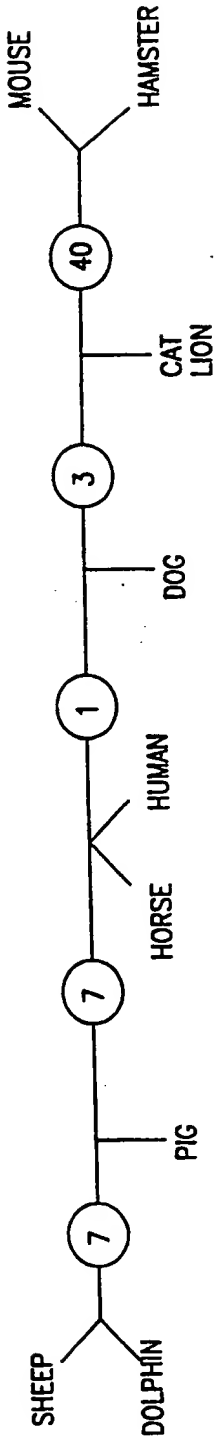

MOUSE
HAMSTER
40
CAT
LION
3
DOG
1
HUMAN
HORSE
7
PIG
7
SHEEP
DOLPHIN

FIG.8

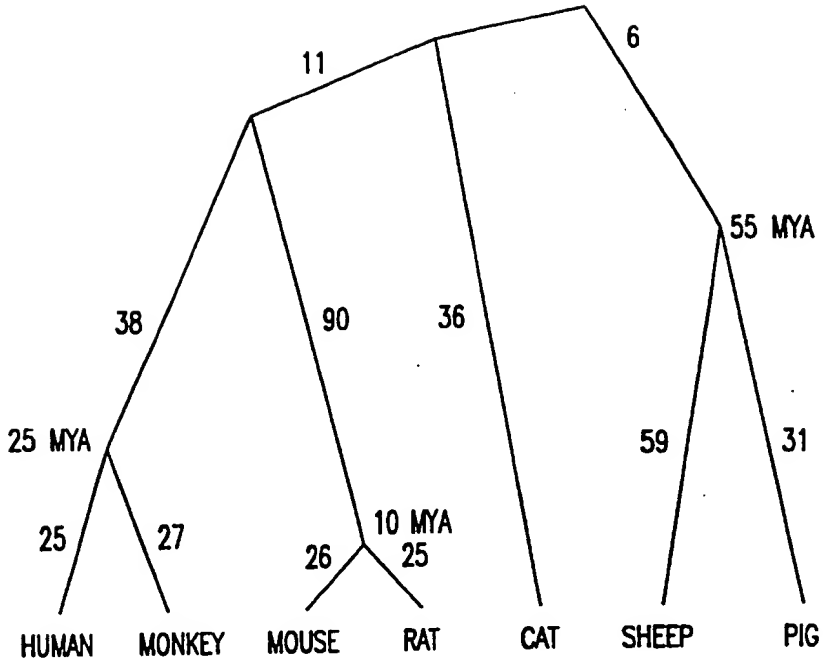

11
6
55 MYA
38
90
36
25 MYA
59
31
25
27
10 MYA
26
25
HUMAN
MONKEY
MOUSE
RAT
CAT
SHEEP
PIG

FIG.7A

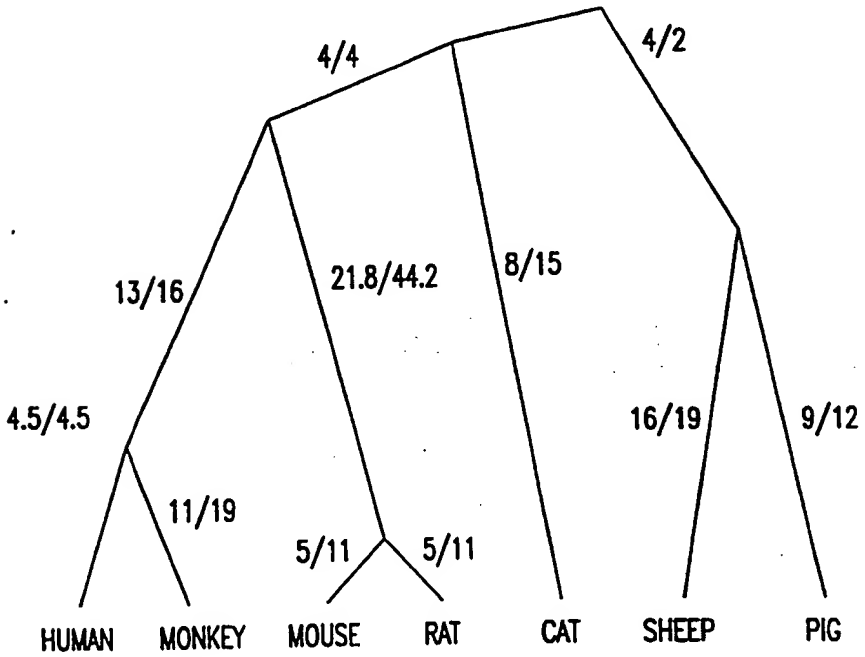

4/4
4/2
13/16
21.8/44.2
8/15
4.5/4.5
16/19
9/12
11/19
5/11
5/11
HUMAN
MONKEY
MOUSE
RAT
CAT
SHEEP
PIG

FIG.7B

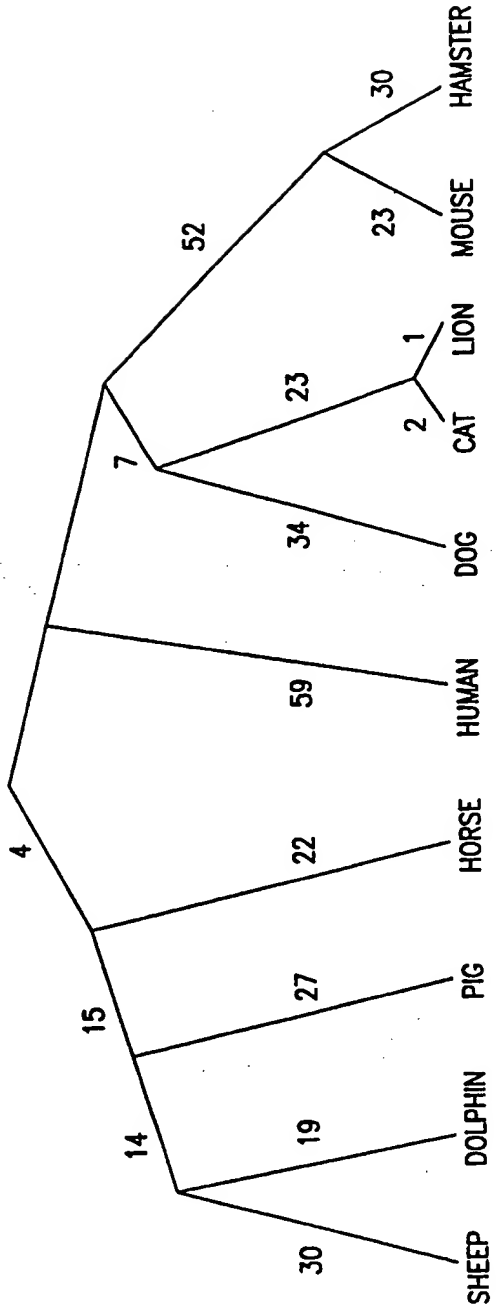

HAMSTER
MOUSE
LION
CAT
DOG
HUMAN
HORSE
PIG
DOLPHIN
SHEEP
30
23
1
2
34
59
22
27
19
30
52
23
7
4
15
14

FIG.8A

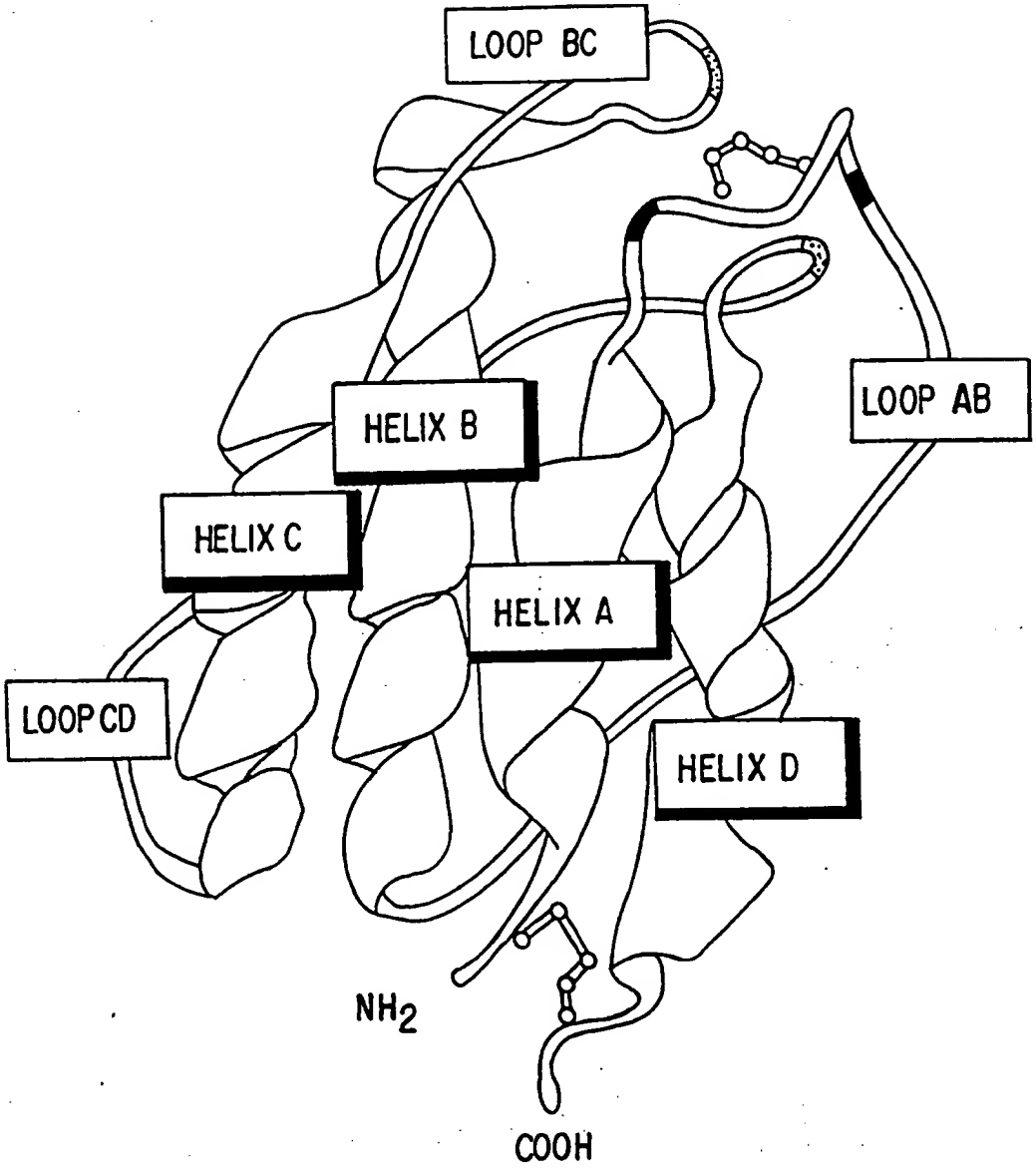

LOOP BC
HELIX B
LOOP AB
HELIX C
HELIX A
LOOP CD
HELIX D
NH_2
COOH
N-GLYCOSYLATION SITES
O-GLYCOSYLATION SITE

FIG.9

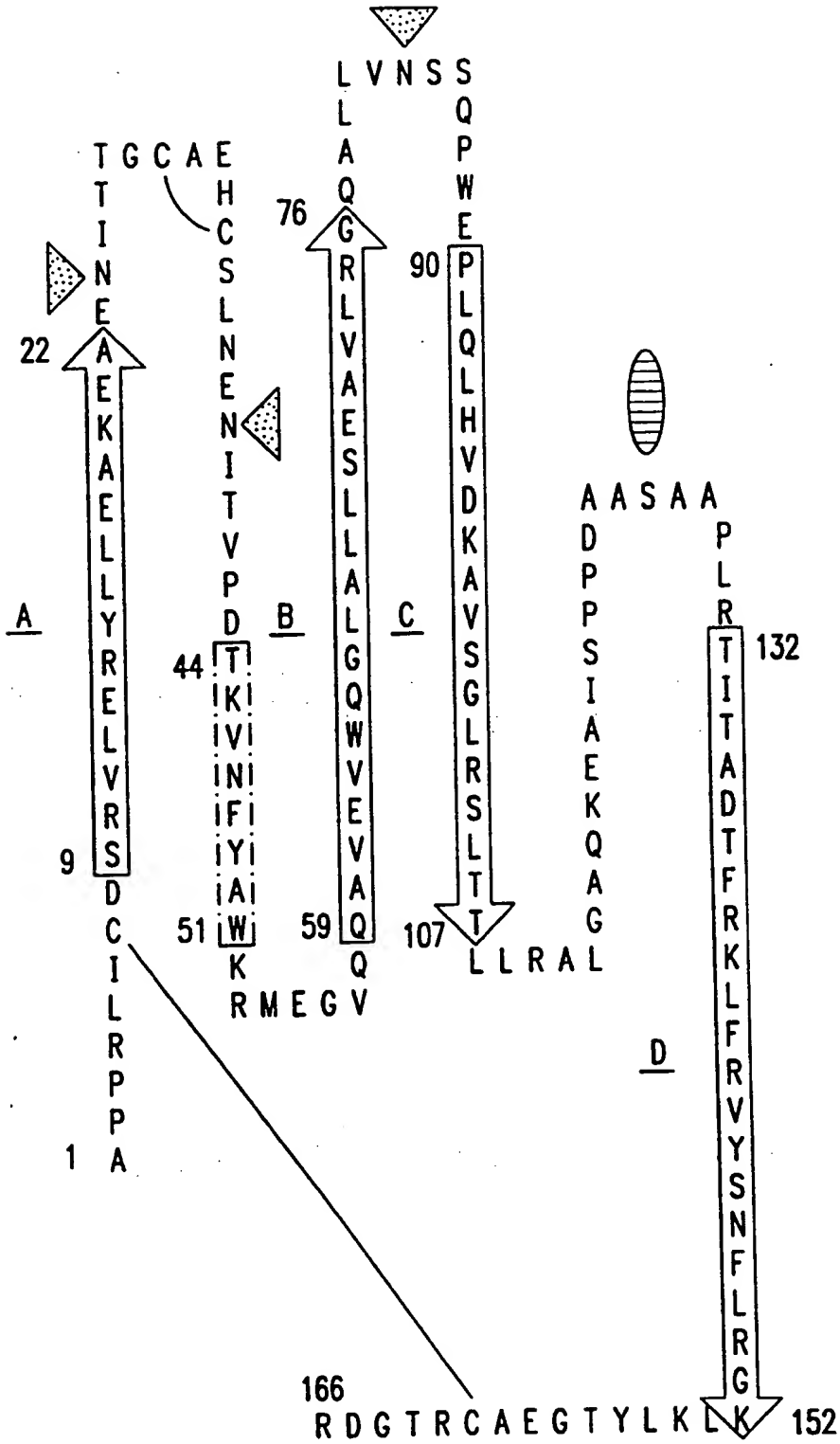

A
B
C
D
1
9
22
44
51
59
76
90
107
132
152
166
T G C A E
T T I N E A E K A E L L Y R E L V R S D C I L R P P A
H C S L N E N I T V P D T I K V I N I F Y I A I W K R M E G V
L V N S S
L A Q G R L V A E S L L A L G Q W V E V A Q Q
Q P W E
P L Q L H V D K A V S G L R S L T T L L R A L
A A S A A
A D P P S I A E K Q A G
P L R
T I T A D T F R K L F R V Y S N F L R G K
R D G T R C A E G T Y L K L K

FIG.9A

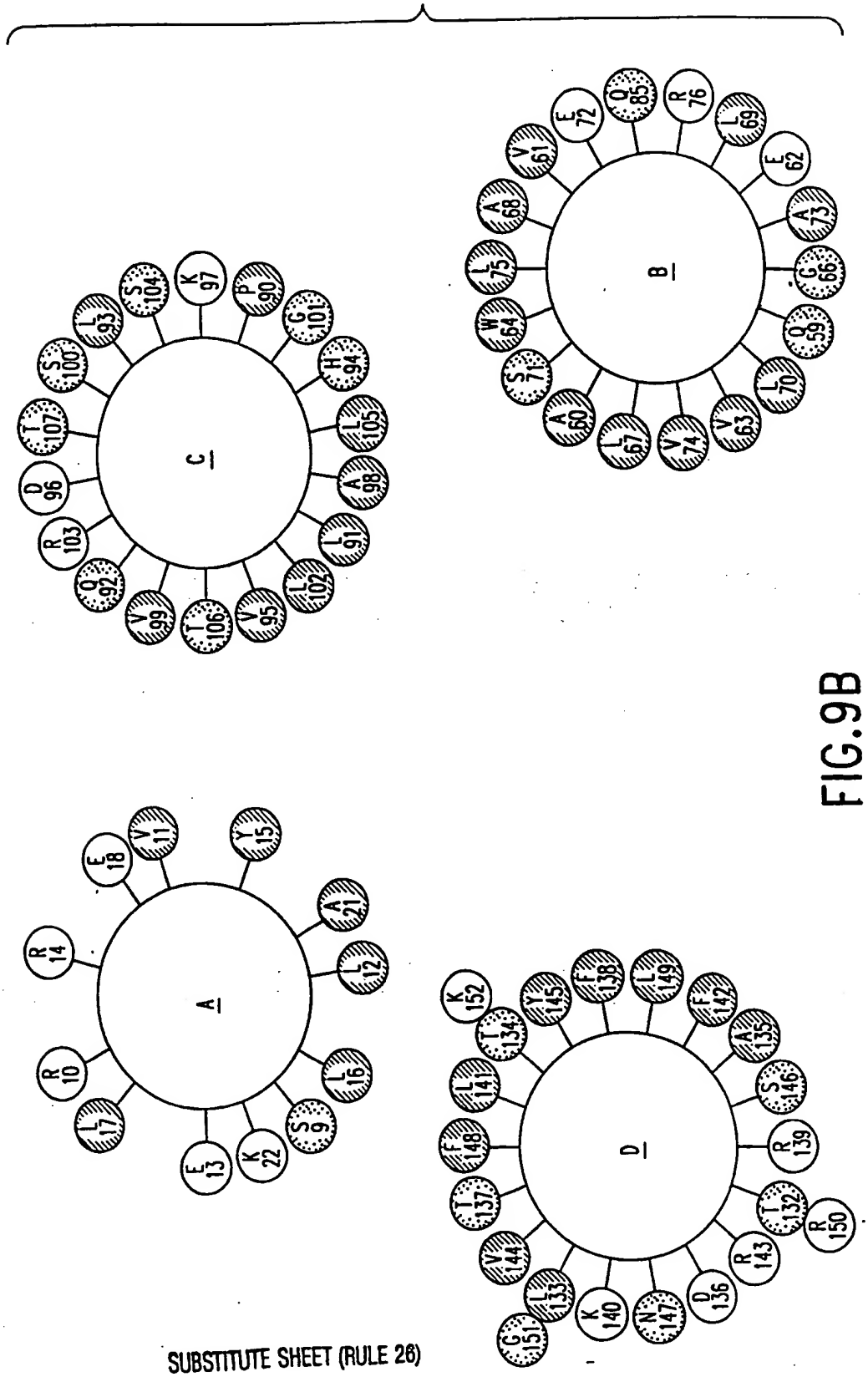

A
R 14
E 18
V 11
Y 15
A 21
L 12
L 16
S 9
K 22
E 13
L 17
R 10
B
Q 85
R 76
L 69
E 62
A 73
G 66
Q 59
L 70
V 63
V 74
L 67
A 60
S 71
W 64
L 75
A 68
V 61
E 72
C
K 97
P 90
G 101
H 94
L 105
A 98
L 91
L 102
V 95
T 106
V 99
Q 92
R 103
D 96
T 107
S 100
L 93
S 104
D
K 152
T 134
Y 145
F 138
L 149
F 142
A 135
S 146
R 139
T 132
R 150
R 143
D 136
N 147
K 140
L 133
V 144
T 137
F 148
L 141

FIG.9B

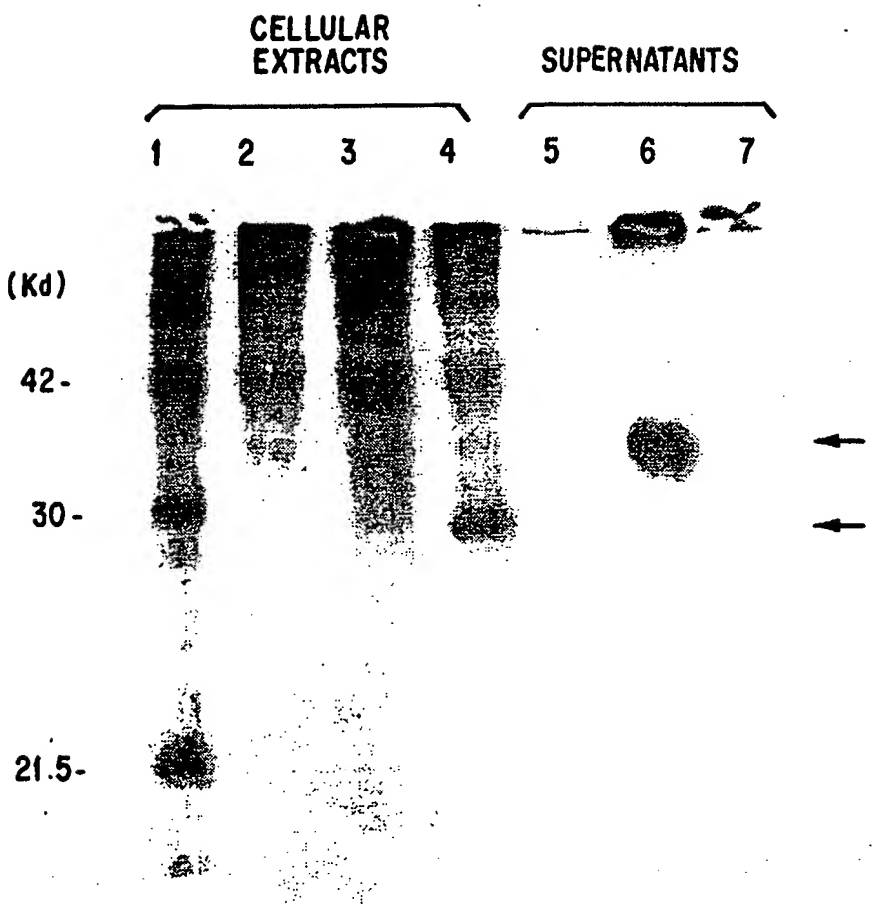
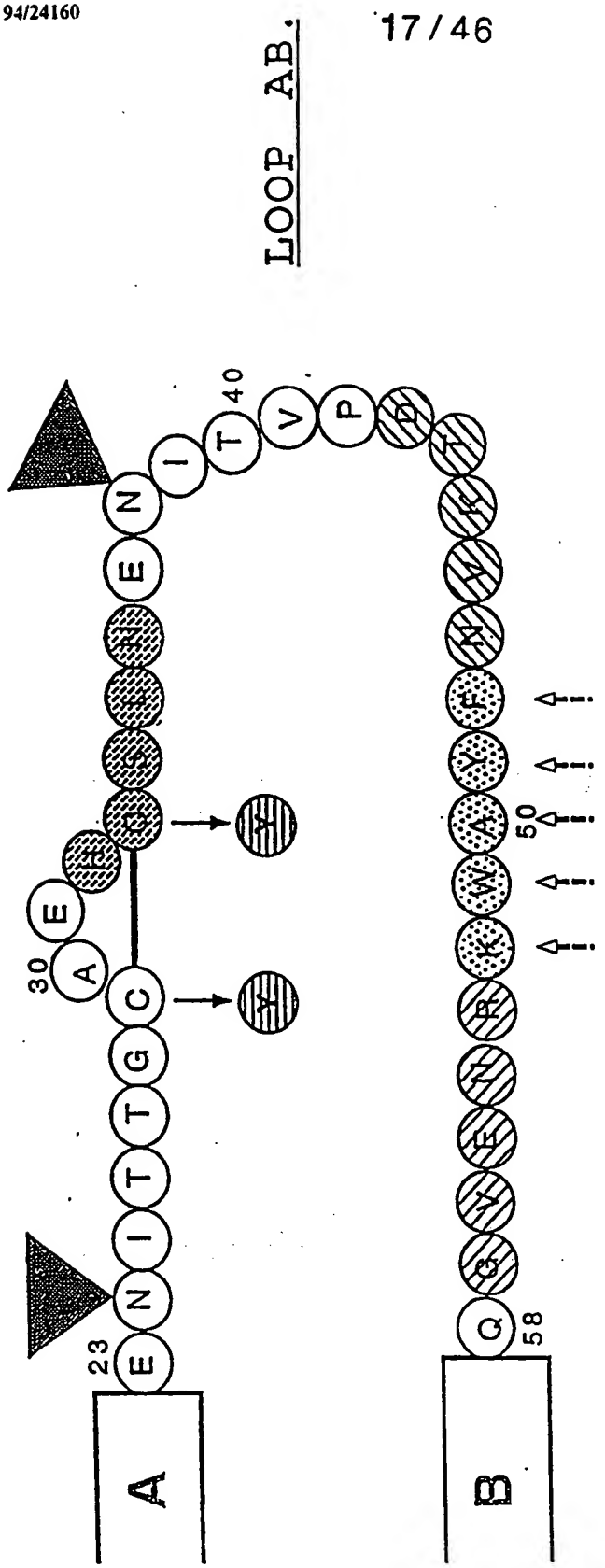


FIG. 10



- FIGURE 11 -

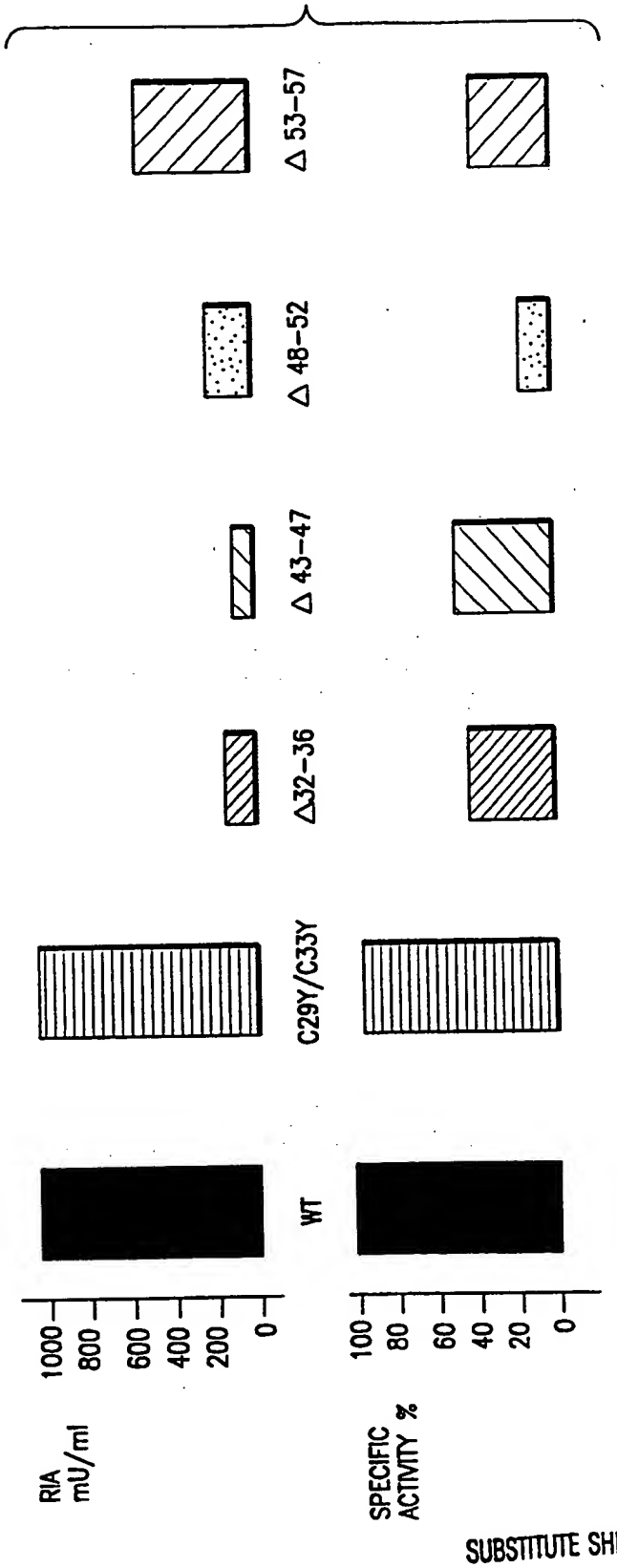
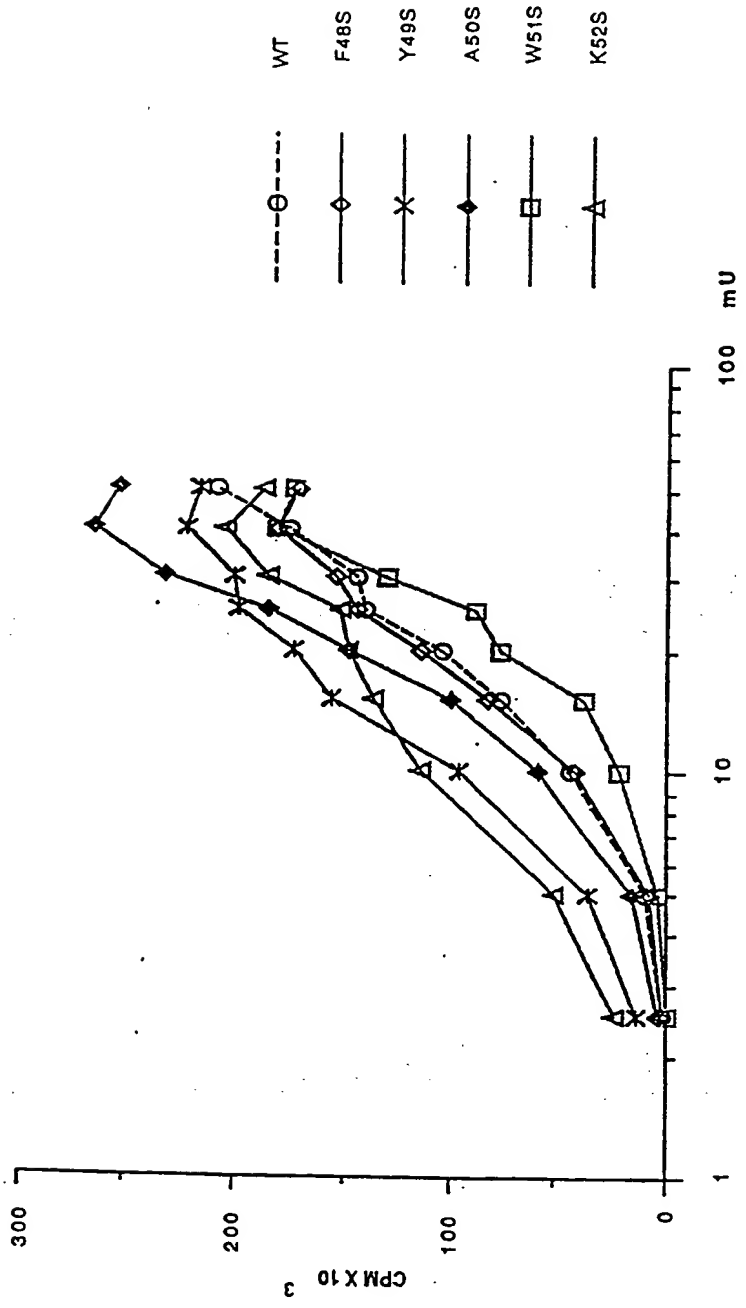
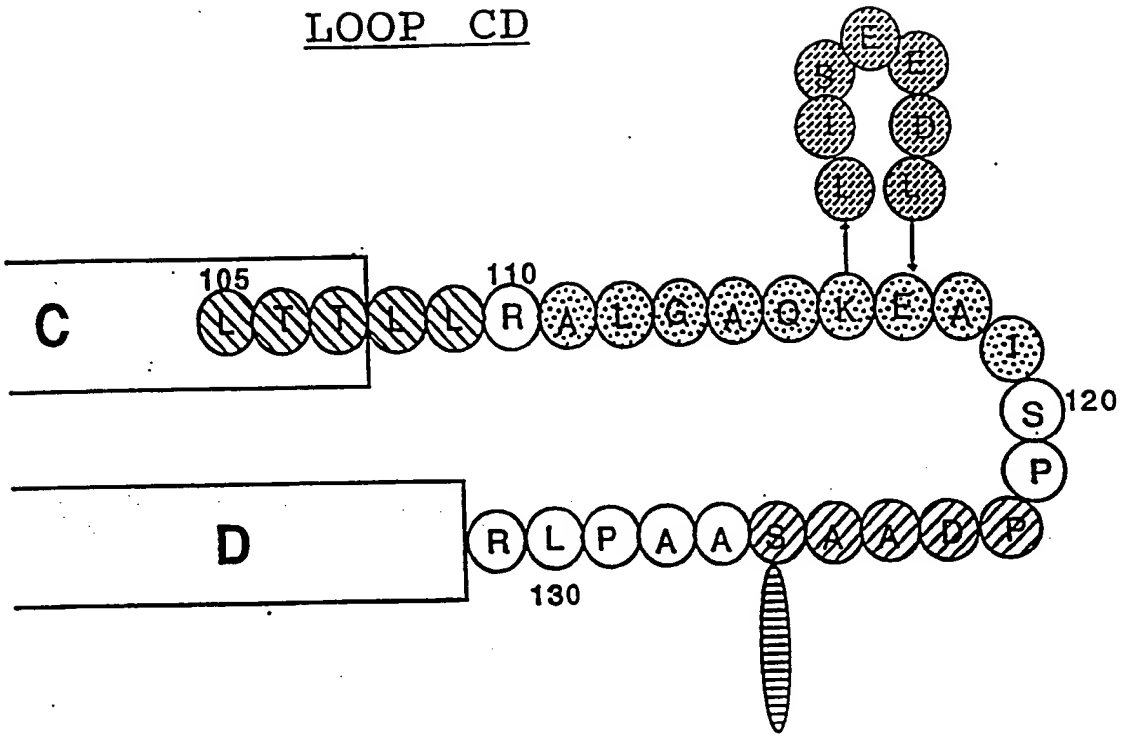

Δ 53-57
Δ 48-52
Δ 43-47
Δ32-36
C29Y/C33Y
WT
1000
800
600
400
200
0
RIA mU/ml
100
80
60
40
20
0
SPECIFIC ACTIVITY %

FIG.11A



- FIGURE 11B -



- FIGURE 12 -

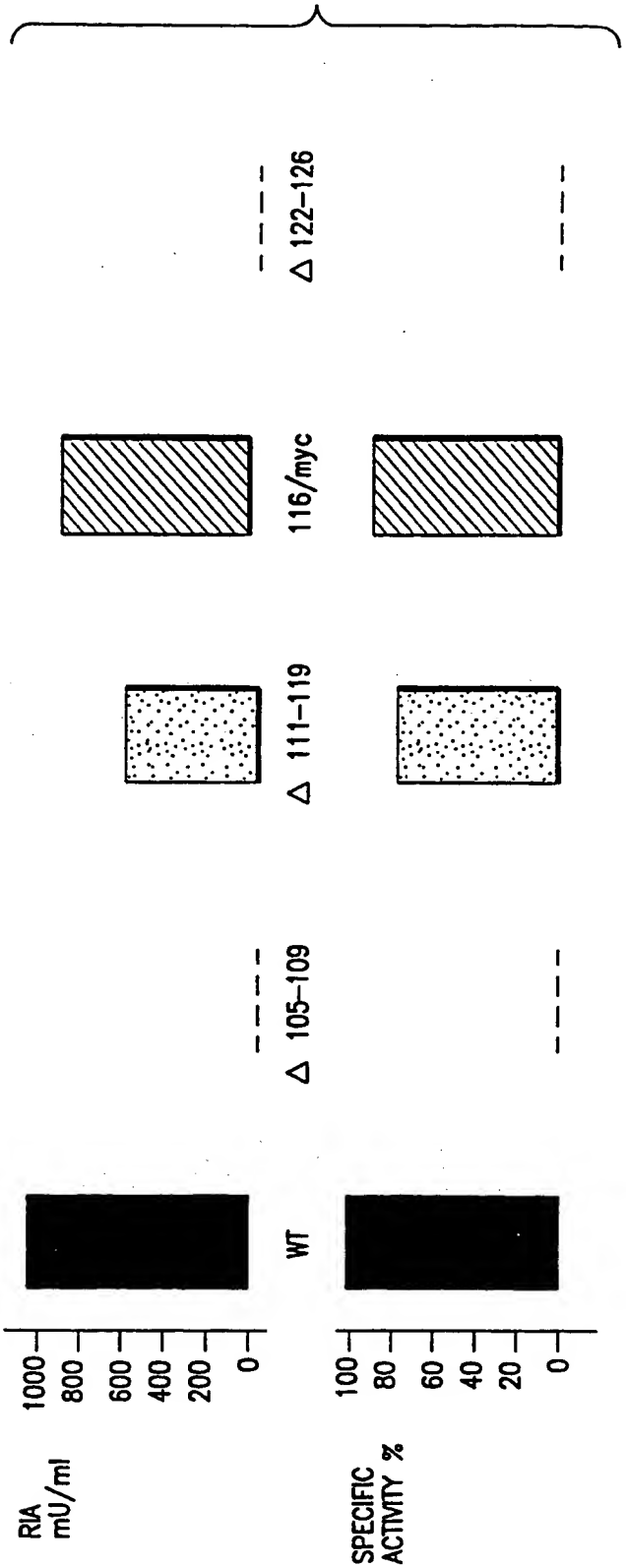

Δ 122–126
116/myc
Δ 111–119
Δ 105–109
WT
RIA mU/ml
1000
800
600
400
200
0
SPECIFIC ACTIVITY %
100
80
60
40
20
0

FIG.12A

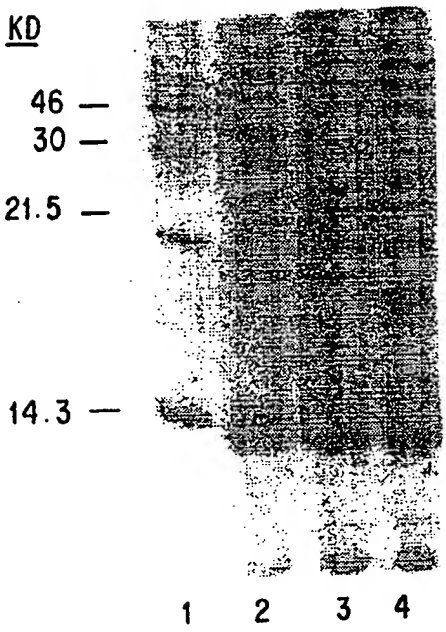
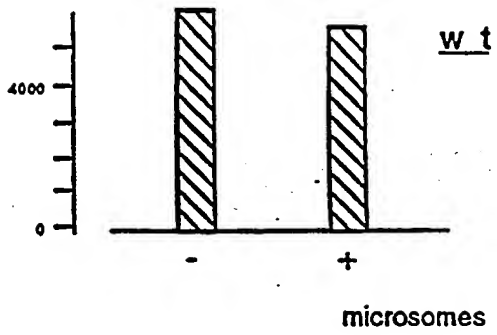
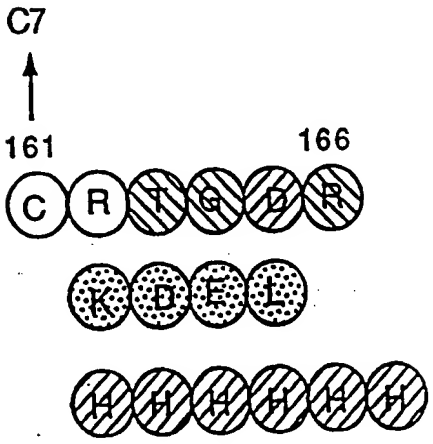


FIG.13



- FIGURE 13A -

C terminus.



- FIGURE 14 -

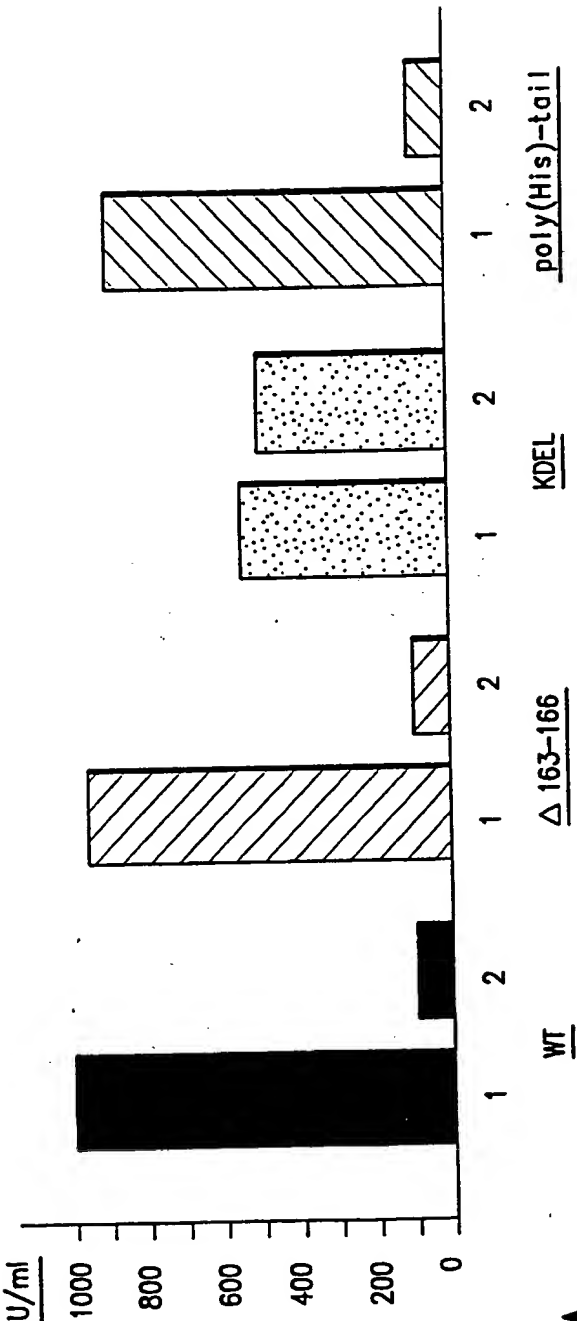


FIG. 14A

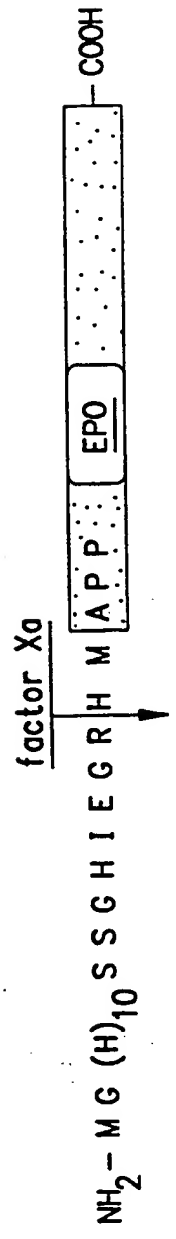


FIG. 15

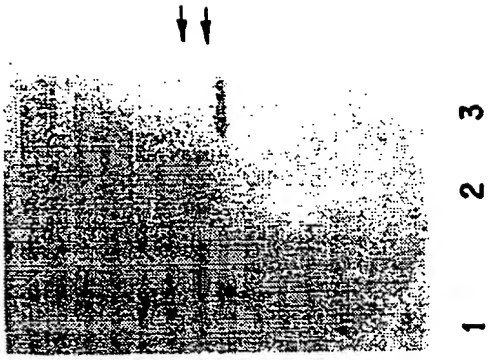


FIG. 15C

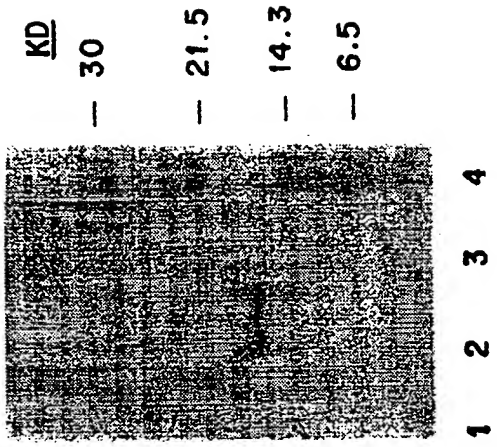


FIG. 15B

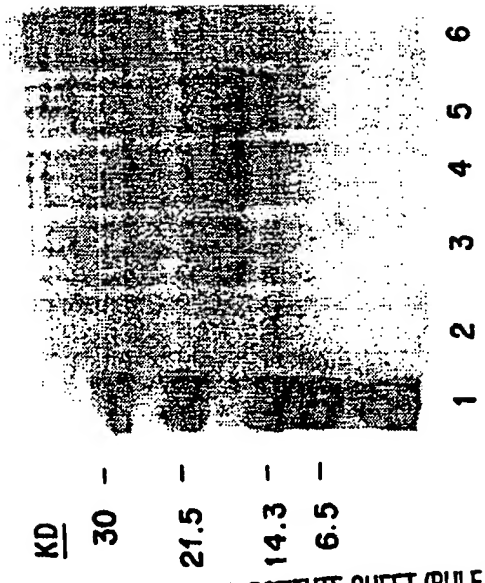


FIG. 15A

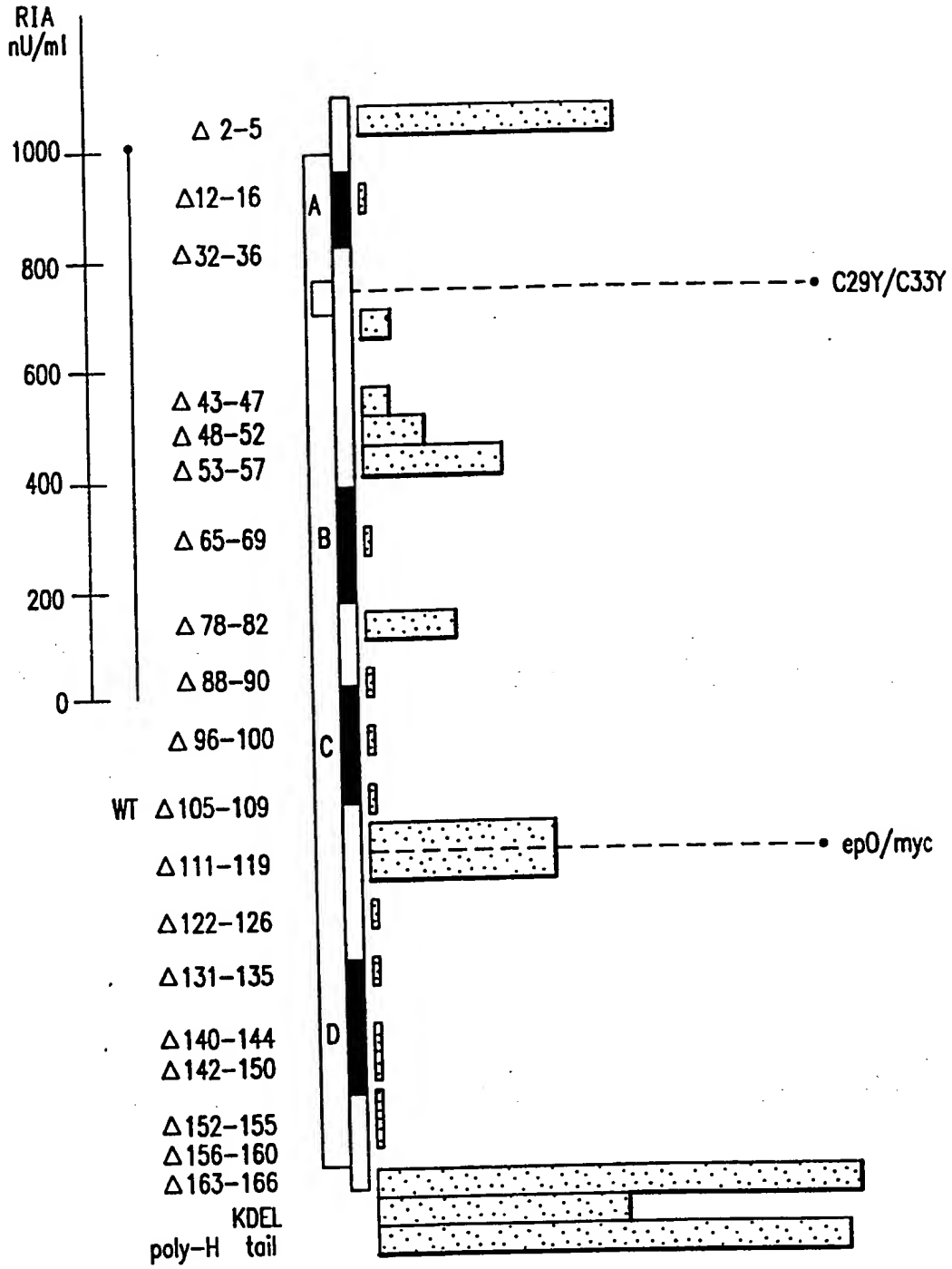
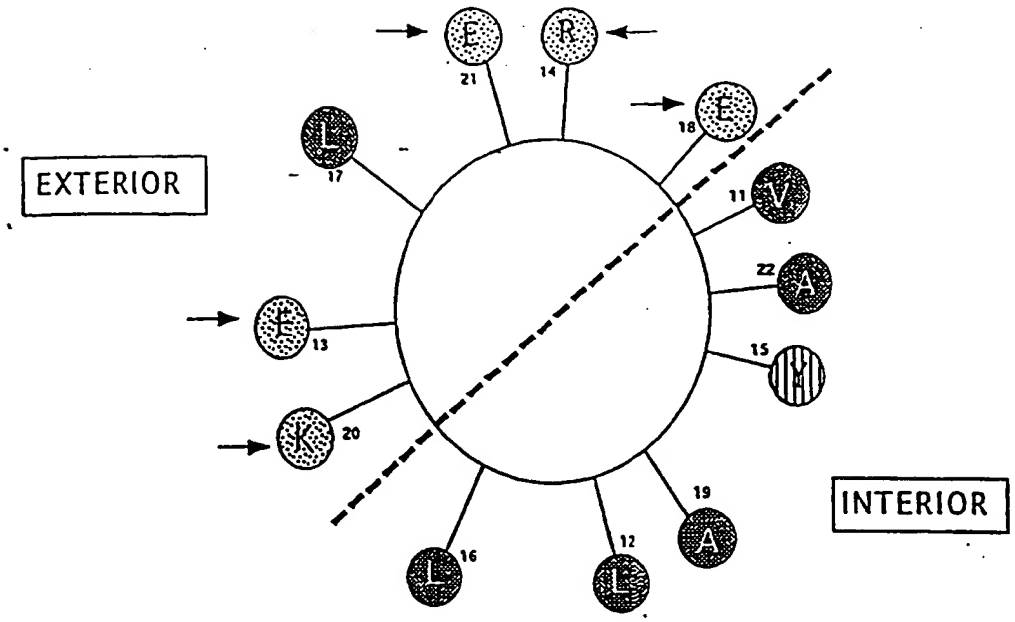
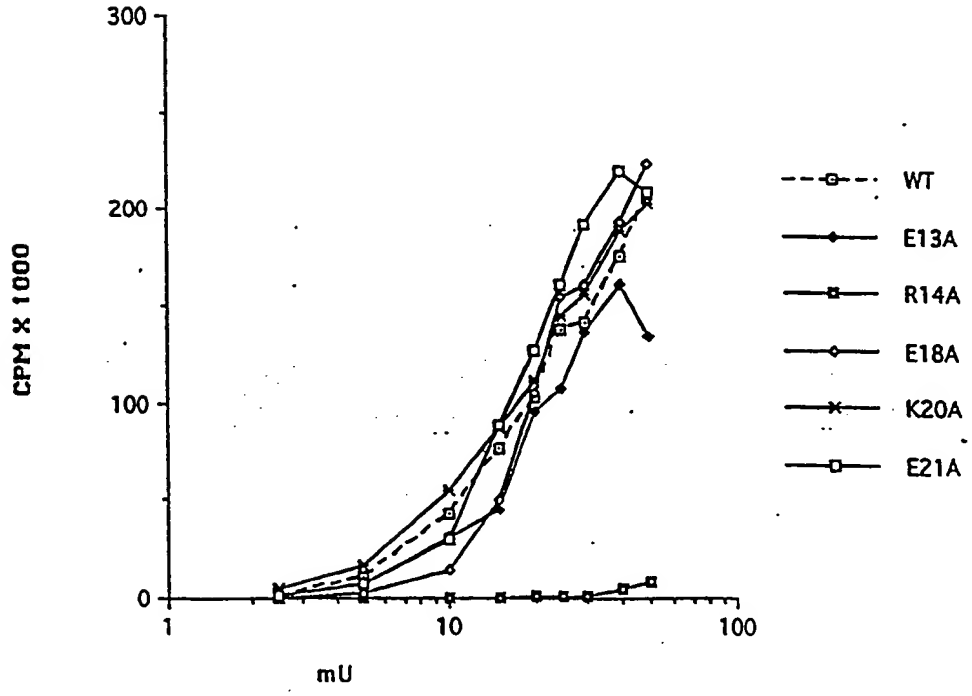

RIA
nU/ml
1000
800
600
400
200
0
Δ 2–5
Δ12–16
Δ32–36
Δ 43–47
Δ 48–52
Δ 53–57
Δ 65–69
Δ 78–82
Δ 88–90
Δ 96–100
WT Δ105–109
Δ111–119
Δ122–126
Δ131–135
Δ140–144
Δ142–150
Δ152–155
Δ156–160
Δ163–166
KDEL
poly-H tail
A
B
C
D
C29Y/C33Y
epO/myc

FIG.16

A HELIX

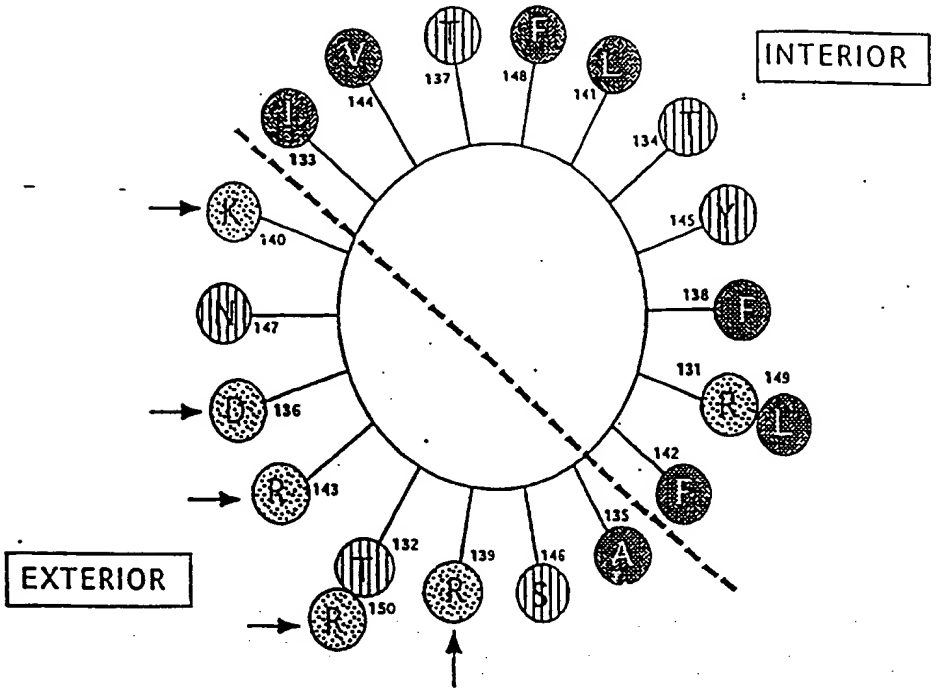


HCD 57

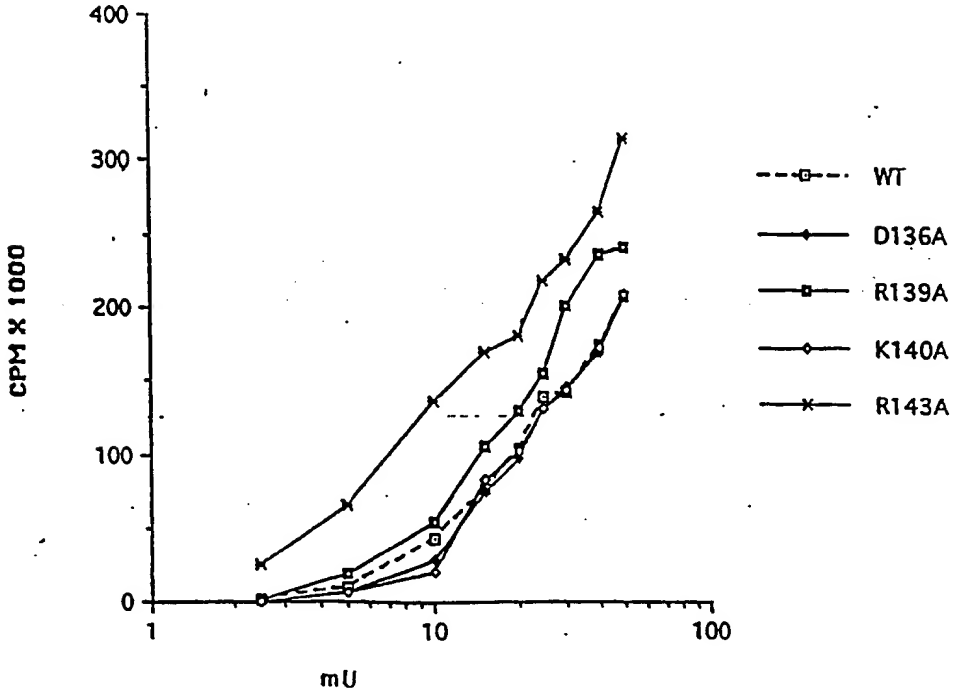


- FIGURE 17 -

D HELIX



HCD 57



- FIGURE 18 -

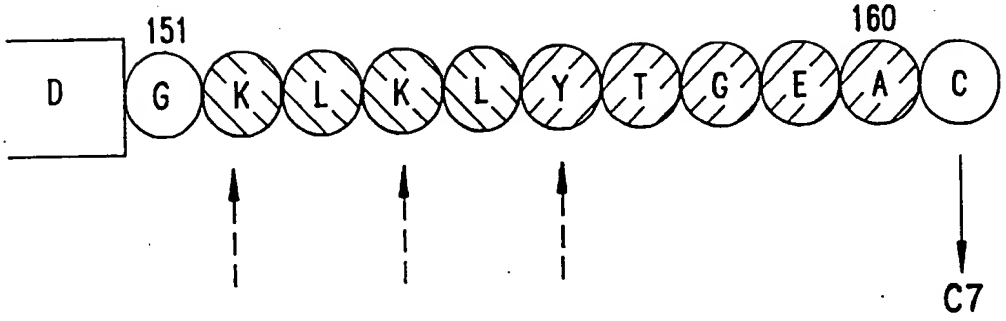


FIG. 19

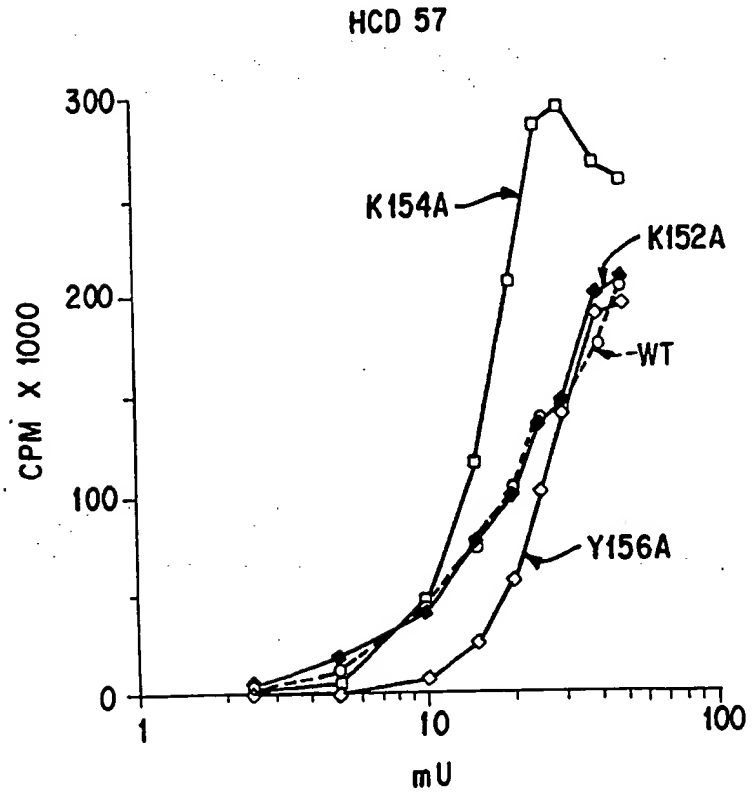


FIG. 19A

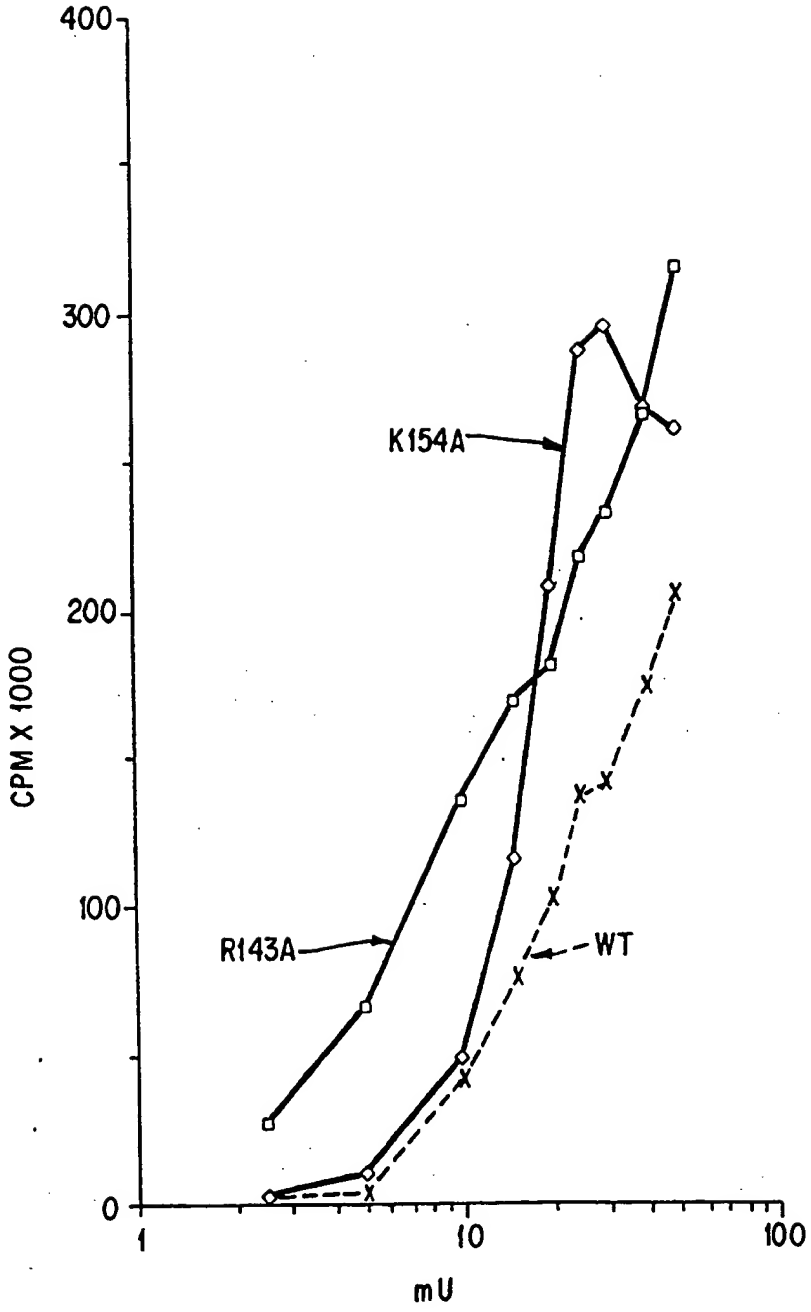


FIG. 20

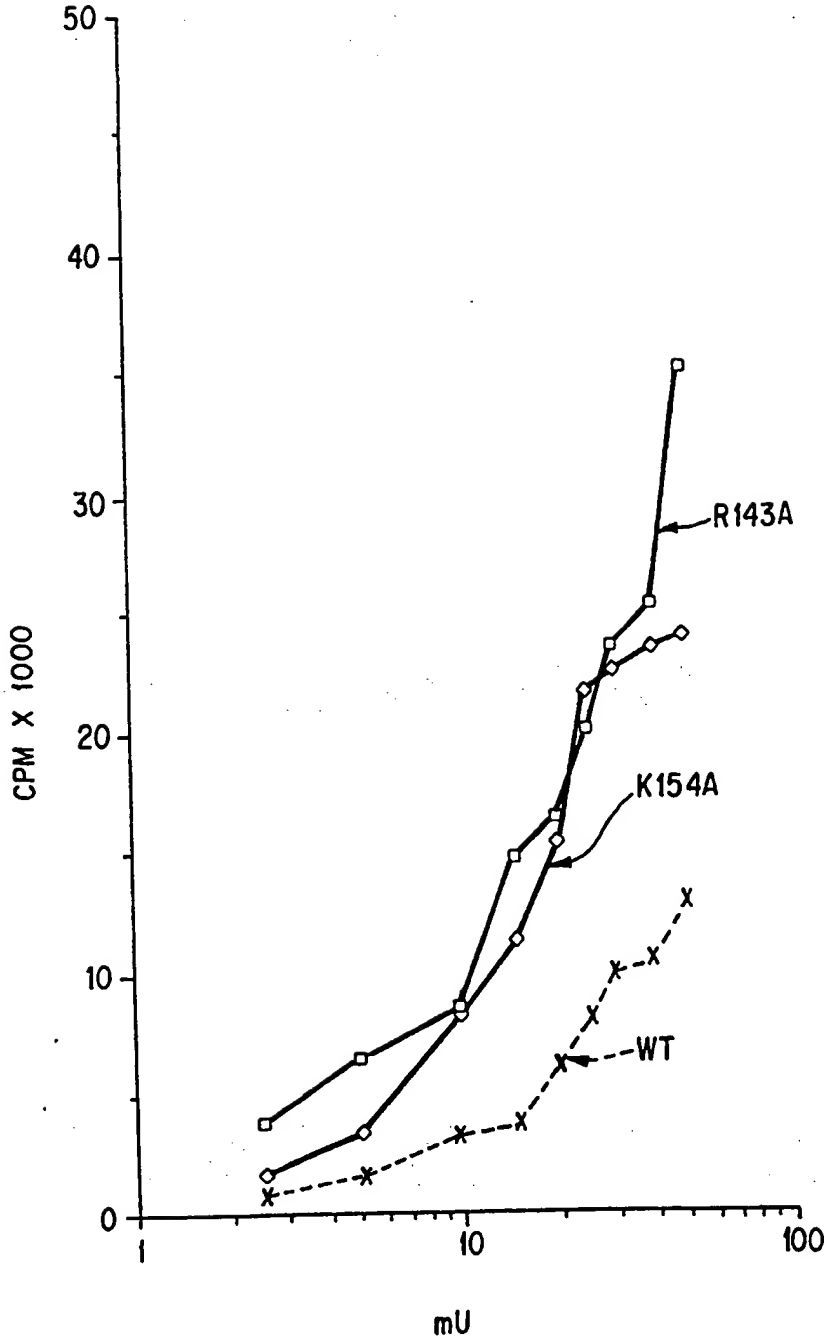


FIG. 21

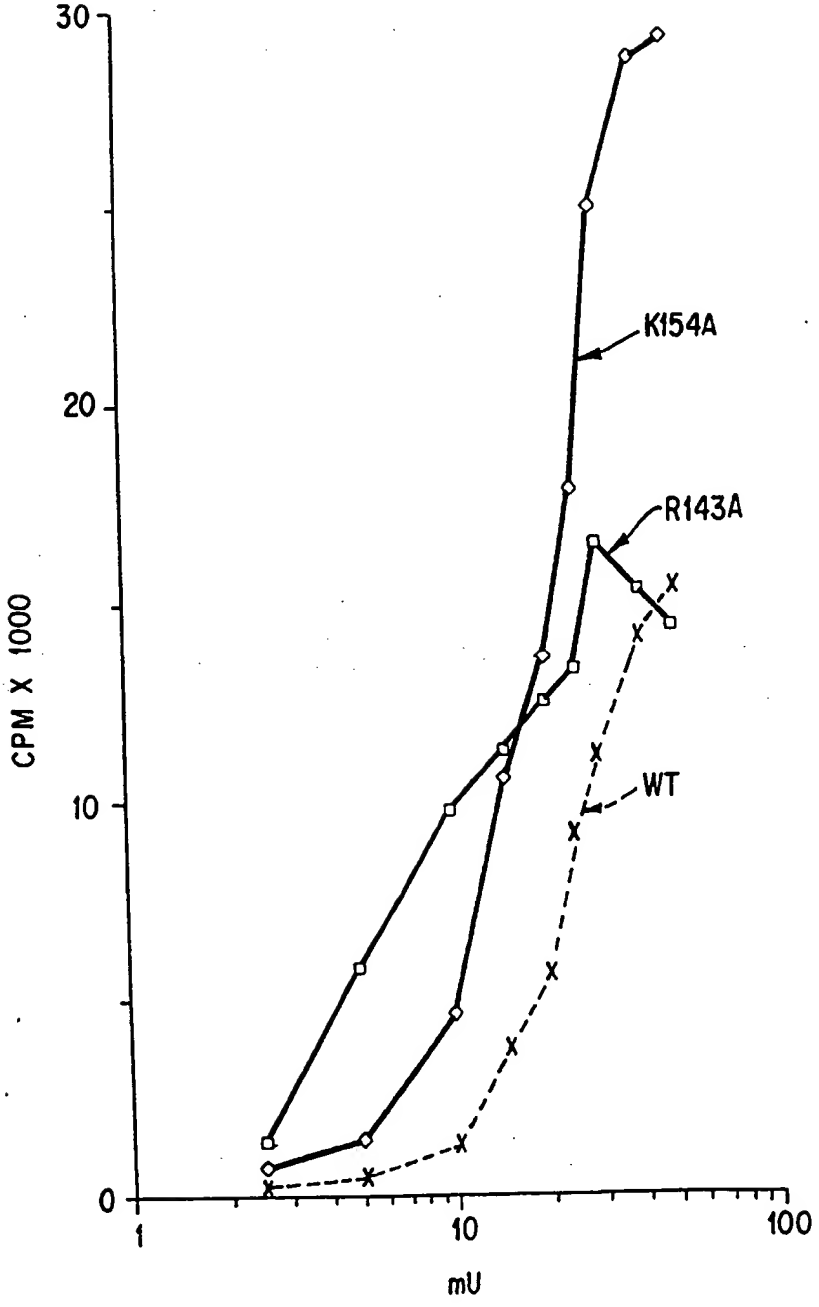


FIG. 22

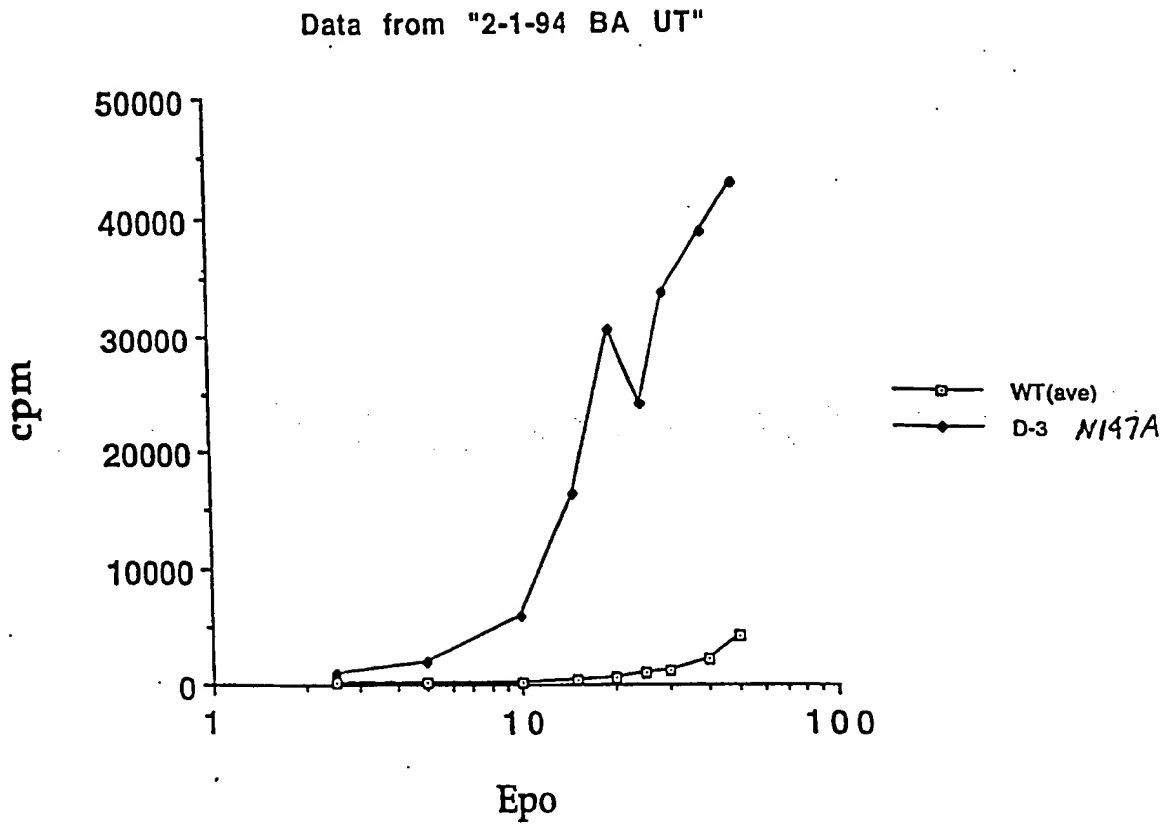

Data from "2-1-94 BA UT"
50000
40000
30000
20000
10000
0
cpm
1
10
100
Epo
WT(ave)
D-3 N147A

FIGURE 23

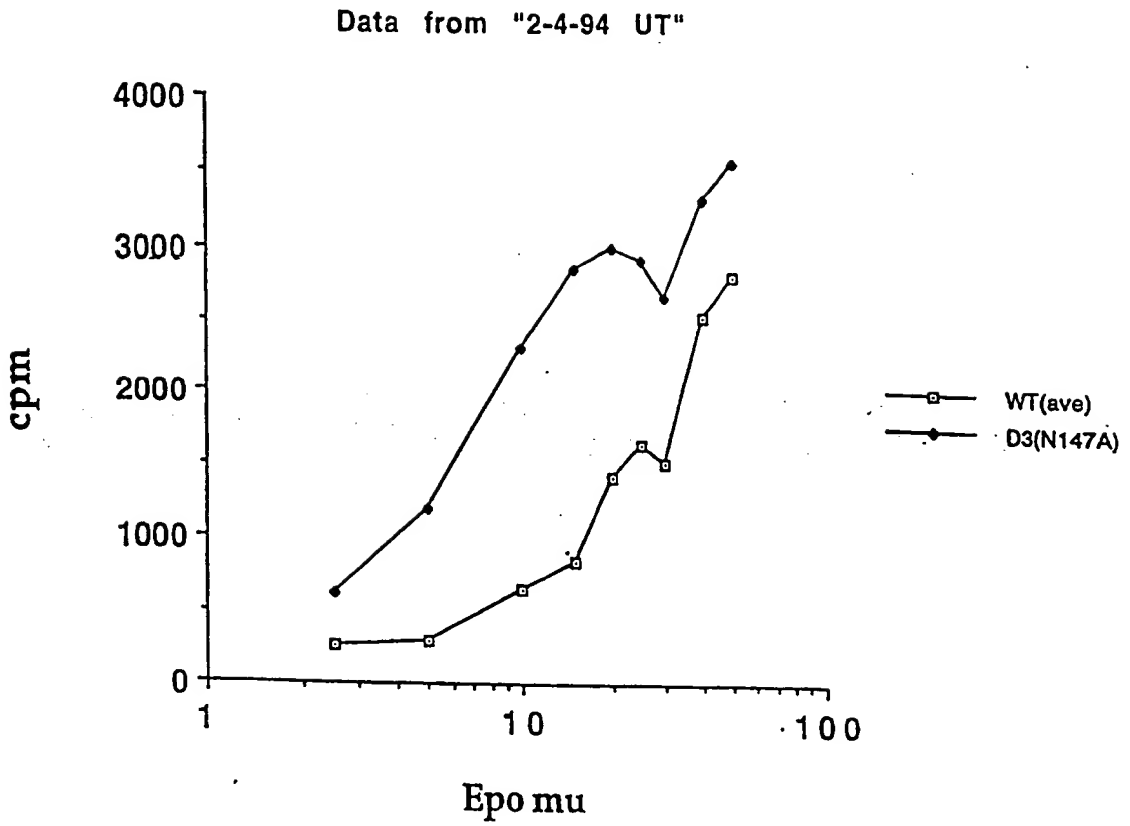

Data from "2-4-94 UT"
4000
3000
2000
1000
0
cpm
1
10
100
Epo mu
WT(ave)
D3(N147A)

FIGURE 24

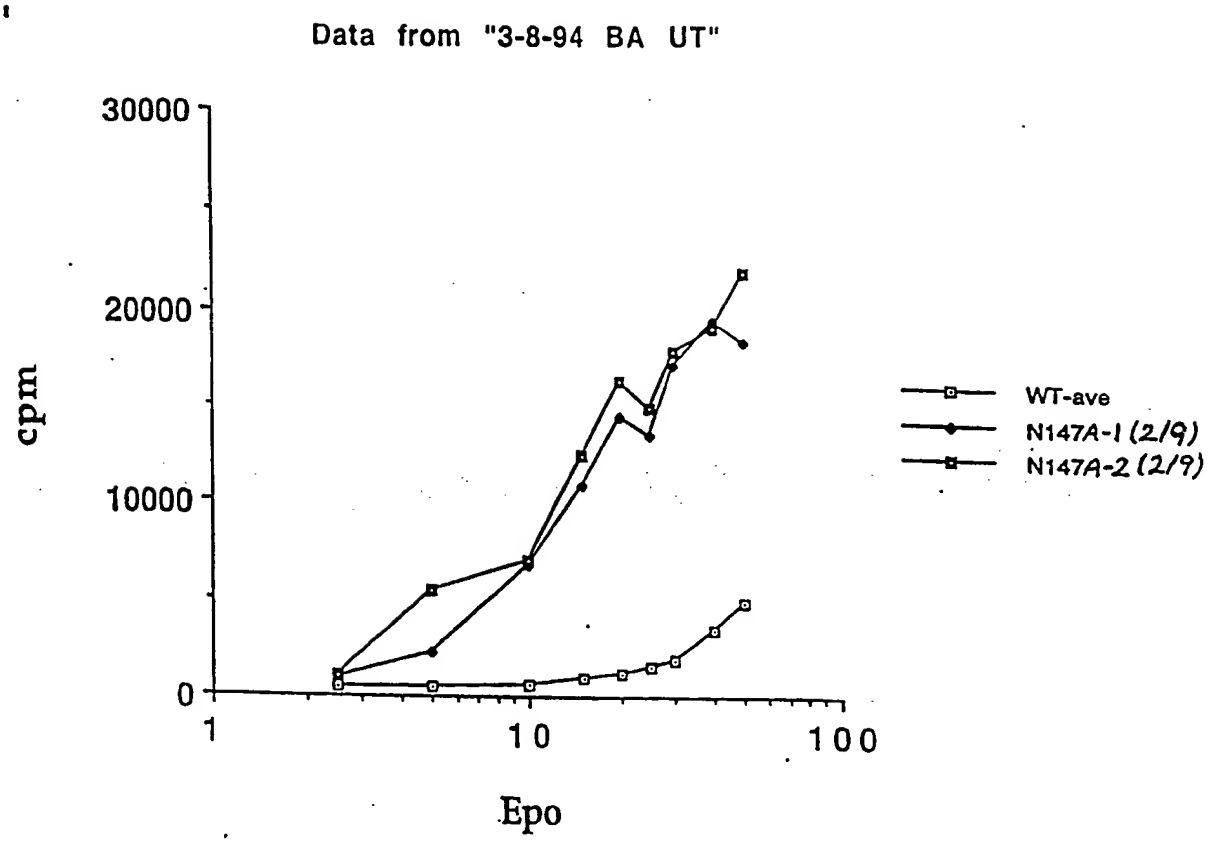

Data from "3-8-94 BA UT"
30000
20000
10000
0
cpm
1
10
100
Epo
WT-ave
N147A-1 (2/9)
N147A-2 (2/9)

FIGURE 25

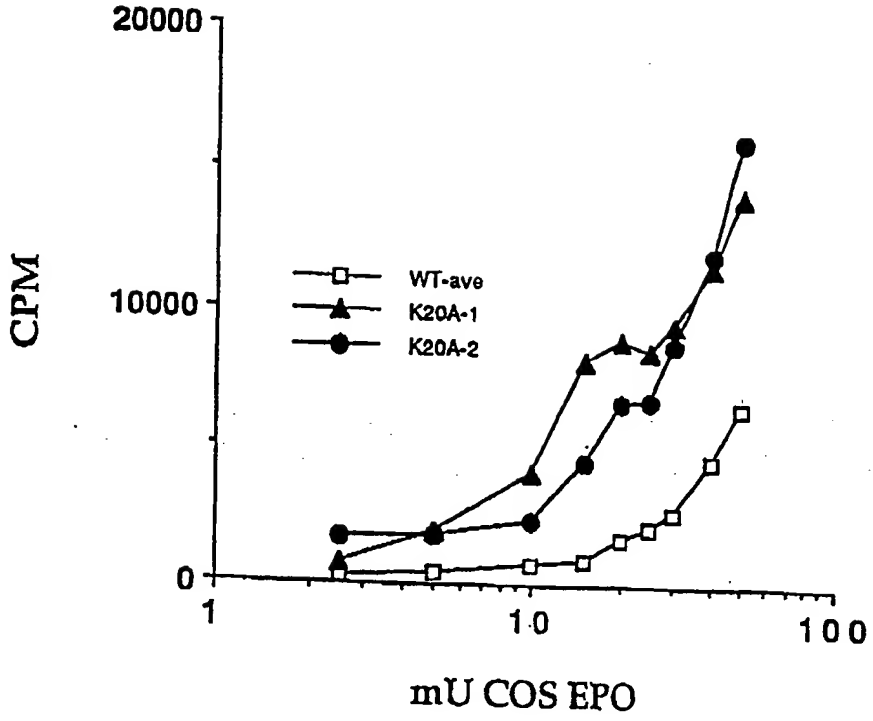


FIGURE 26

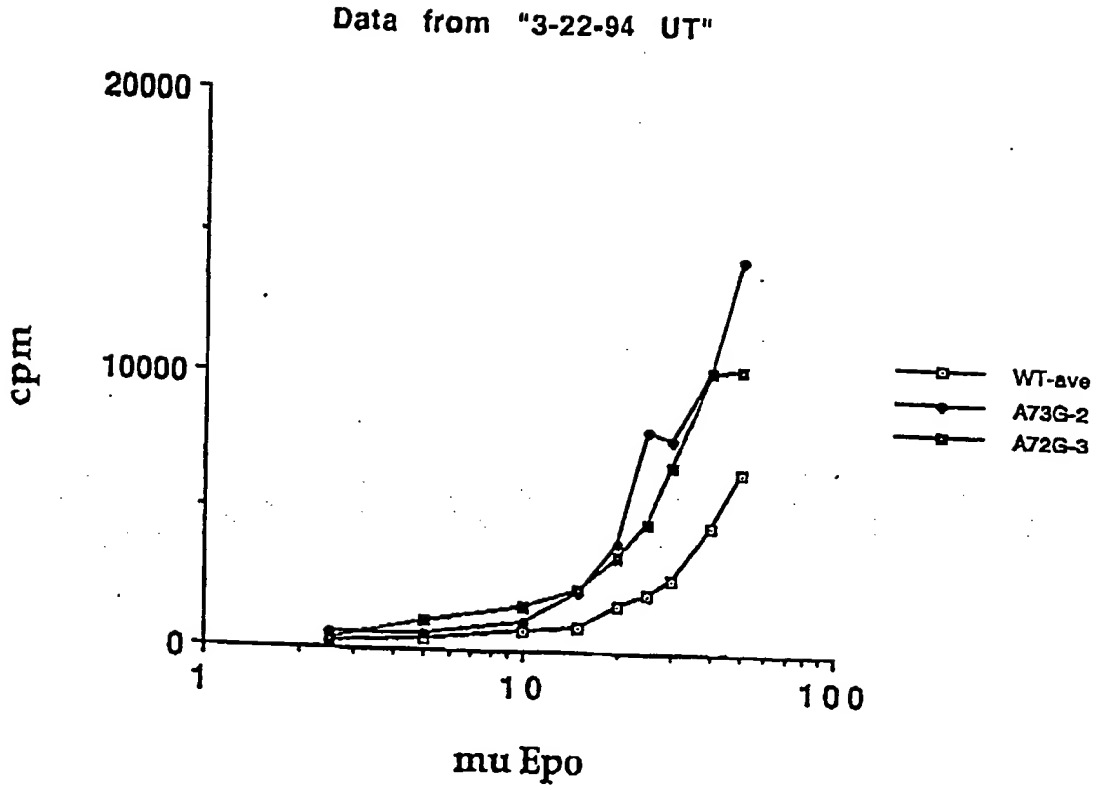
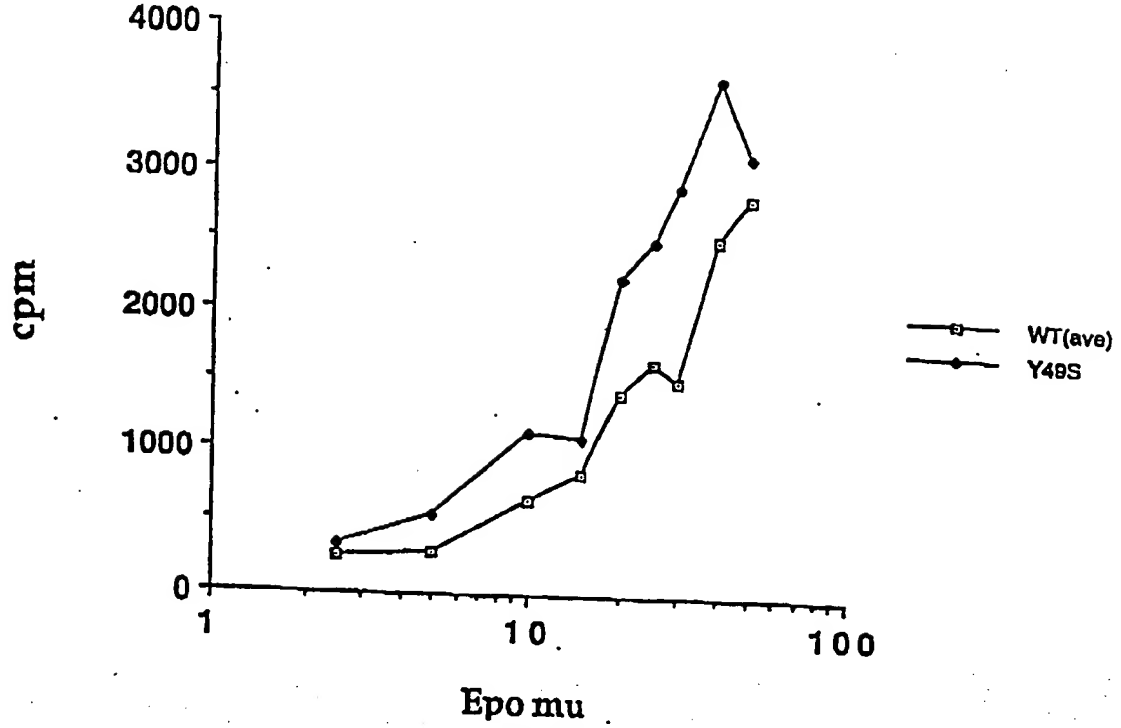

Data from "3-22-94 UT"
20000
10000
0
cpm
1
10
100
mu Epo
WT-ave
A73G-2
A72G-3

FIGURE 27

Data from "2-4-94 UT"



Data from "2-4-94 HCD"

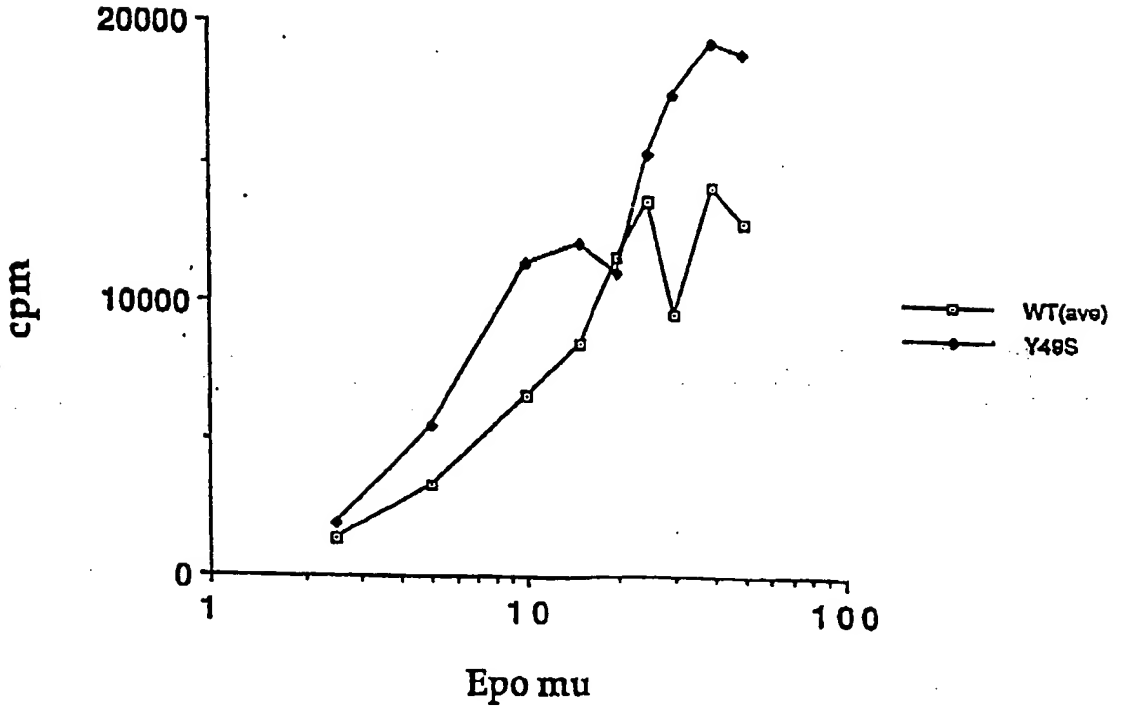


FIGURE 28

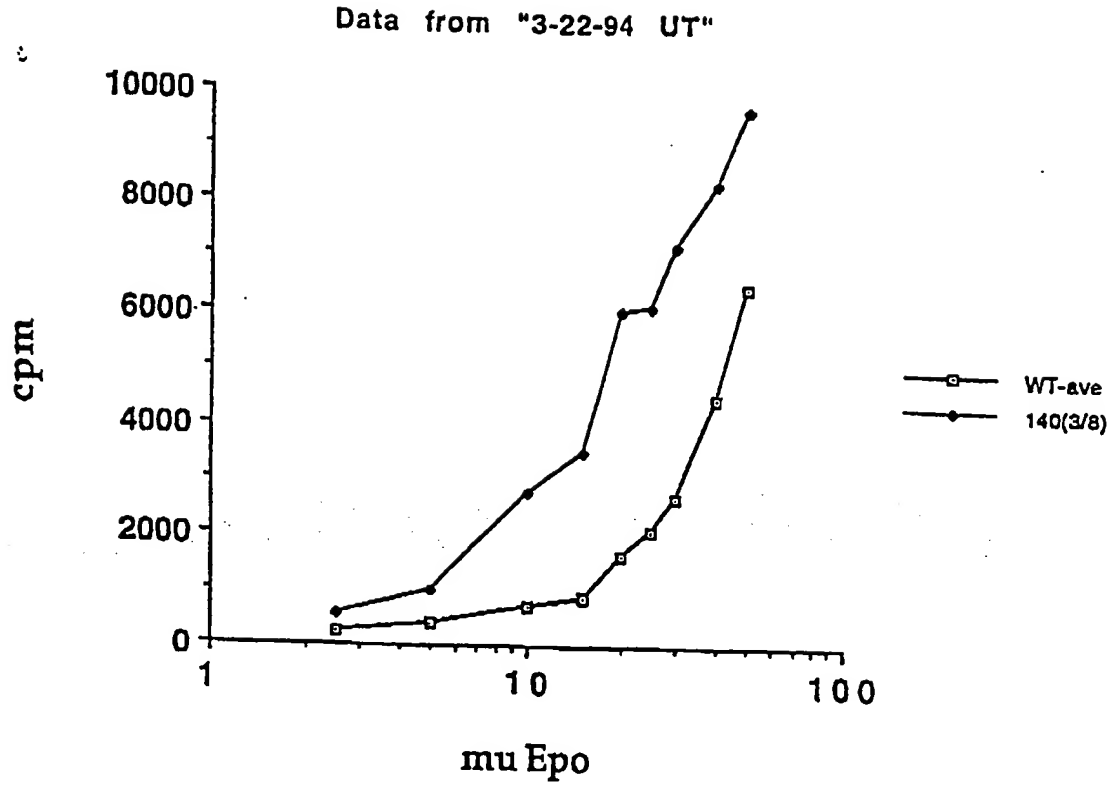

Data from "3-22-94 UT"
10000
8000
6000
4000
2000
0
cpm
1
10
100
mu Epo
WT-ave
140(3/8)

FIGURE 29

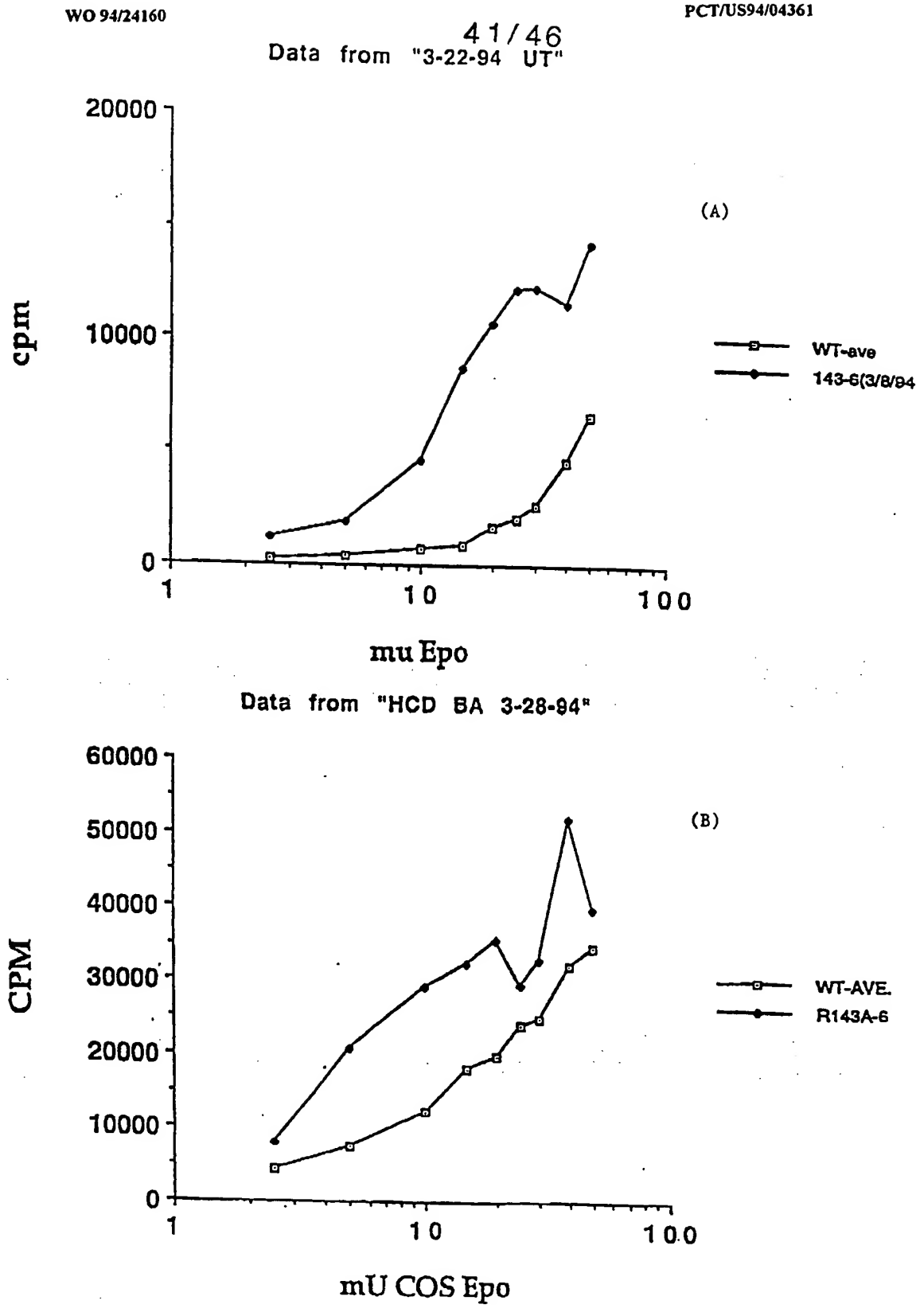

Data from "3-22-94 UT"
20000
10000
0
cpm
1
10
100
mu Epo
(A)
WT-ave
143-6(3/8/94
Data from "HCD BA 3-28-94"
60000
50000
40000
30000
20000
10000
0
CPM
1
10
100
mU COS Epo
(B)
WT-AVE.
R143A-6

FIGURE 30

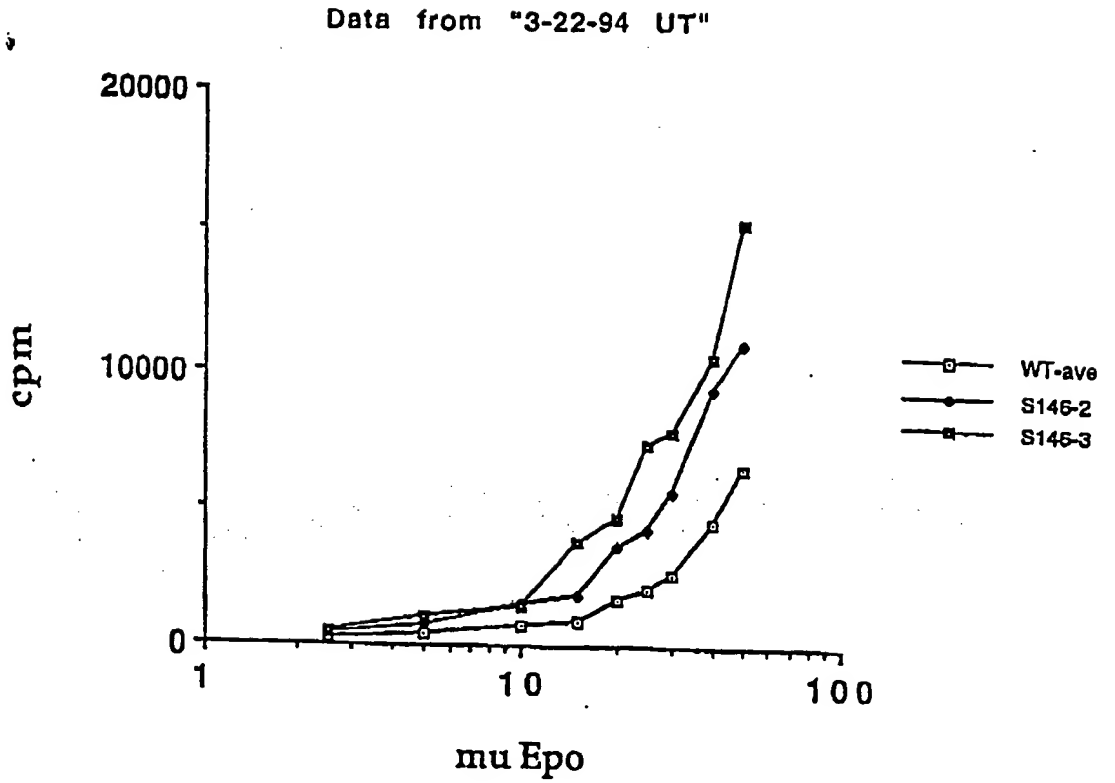

Data from "3-22-94 UT"
20000
10000
0
cpm
1
10
100
mu Epo
WT-ave
S146-2
S146-3

FIGURE 31

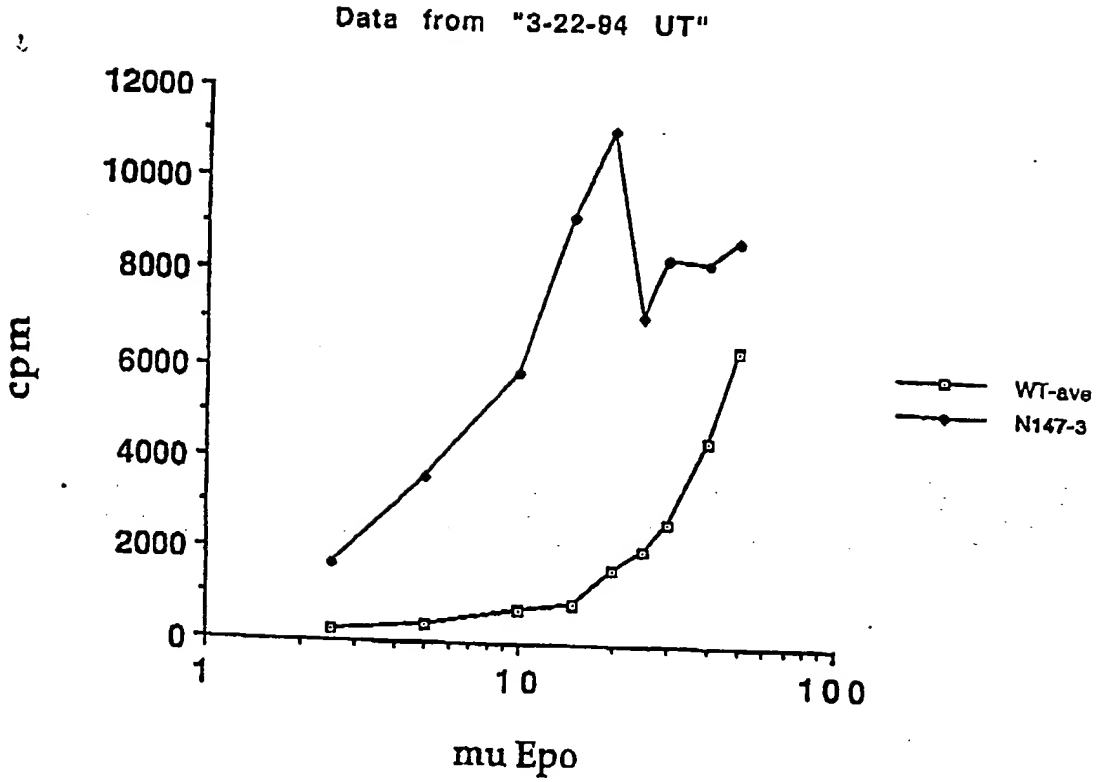

Data from "3-22-94 UT"
12000
10000
8000
6000
4000
2000
0
cpm
1
10
100
mu Epo
WT-ave
N147-3

FIGURE 32

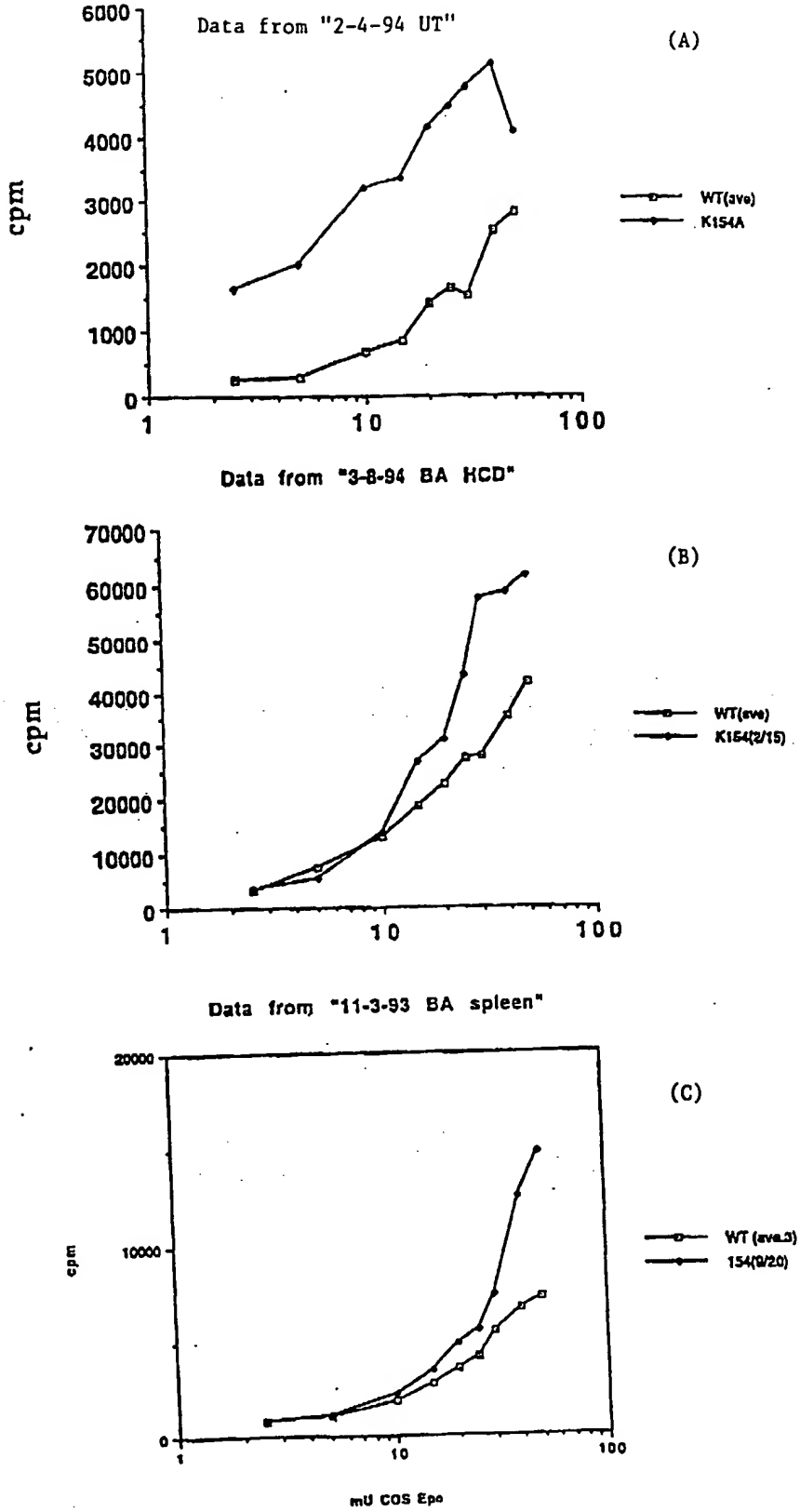

Data from "2-4-94 UT"
(A)
cpm
6000
5000
4000
3000
2000
1000
0
1
10
100
WT(ave)
K154A
Data from "3-8-94 BA HCD"
(B)
cpm
70000
60000
50000
40000
30000
20000
10000
0
1
10
100
WT(ave)
K154(2/15)
Data from "11-3-93 BA spleen"
(C)
cpm
20000
10000
0
1
10
100
mU COS Epo
WT (ave.3)
154(9/20)

FIGURE 33

SPECIFIC BIOACTIVITY (% wt h EPO)				
Helix	Mutein	UT7	HCD	Spleen
A	S 9 A	148 ± 35 [3/3]	123 ± 13 [3/3]	-
A	R 10 A	91 [1/1]	123 [1/1]	-
A	E 13 A	111,74 [2/2]	80, 99 [2/2]	-
A	R 14 A	58,56 [2/2]	2, 30 [2/2]	18, 24 [2/2]
A	R 14 L	91 ± 8 [4/6]	86 ± 15 [4/6]	67, 111 [2/2]
A	R 14 E	11 ± 6 [3/5]	12 ± 1 [5/3]	17, 17 [2/2]
A	L 17 A	110 [1/1]	95 [1/1]	-
A	E 18 A	70 ± 17 [3/3]	100, 90 [2/2]	-
A	K 20 A	252 ± 81 [4/4]	103 ± 21 [4/4]	-
A	E 21 A	99, 69 [2/2]	50 [1/1]	-
A-B	C 29 Y / C 33 Y	75 [1/1]	42 [1/1]	-
A-B	K 45 A	110, 101 [2/2]	100, 100 [2/2]	-
A-B	F 48 S	146, 110 [2/2]	100, 61 [2/2]	-
A-B	Y 49 S	150 ± 7 [3/3]	146, 138 [2/2]	-
A-B	A 50 S	128, 92 [2/2]	140, 37 [2/2]	-
A-B	W 51 S	101, 78 [2/2]	89 ± 7 [3/3]	-
A-B	K 52 S	99, 90 [2/2]	99, 88 [2/2]	-
B	Q 59 A	156, 125 [1/2]	130, 134 [1/2]	-
B	E 62 A	101, 85 [2/2]	63,78 [2/2]	82, 111 [2/2]
B	W 64 A	102, 112 [2/2]	87,143 [2/2]	90, 107 [2/2]
B	Q 65 A	96 [1/1]	140 [1/1]	99 [1/1]
B	G 66 A	130 [1/1]	103 [1/1]	85 [1/1]
B	L 69 A	94 [1/1]	84 [1/1]	-
B	S 71 A	64 [1/1]	130 [1/1]	85 [1/1]
B	E 72 A	NS	-	-
B	A 73 G	242 ± 25 [2/3]	104 [1/1]	-
B	R 76 A	109 [1/1]	138 [1/1]	-

FIG.34

Helix	Mutein	UT7	HCD	Spleen
C	Q 92 A	112 [1/1]	98, 95 [2/2]	113, 91 [2/2]
C	L 93 A	123, 126 [2/2]	127, 95 [2/2]	-
C	H 94 A	NS	-	-
C	D 96 A	NS	-	-
C	K 97 A	84 ± 9 [2/4]	77 ± 27 [2/3]	94 [1/1]
C	S 100 A	98, 85 [2/2]	131 [1/1]	104 [1/1]
C	G 101 A	166 ± 52 [3/3]	146 ± 26 [3/4]	73, 152 [2/2]
C	R 103 A	3 ± 3 [3/5]	7 ± 1 [2/3]	6 [1/1]
C	S 104 A	114 ± 28 [2/4]	35 ± 17 [2/3]	-
C	T 106 A	120 [1/1]	108 [1/1]	-
C-D	L 108 A	158 ± 41 [2/4]	48 ± 6 [2/4]	12, 52 [2/2]
D	D 136 A	108, 114 [2/2]	88 ± 20 [3/3]	-
D	R 139 A	171, 91 [2/2]	120, 109 [2/2]	90, 131 [1/2]
D	K 140 A	189 ± 30 [3/3]	110 ± 19 [3/3]	-
D	R 143 A	276 ± 85 [3/3]	150 ± 27 [3/3]	-
D	S 146 A	198, 269 [1/2]	58, 110 [1/2]	-
D	N 147 A	457 ± 185 [3/6]	115 ± 22 [5/6]	-
D	R 150 A	181 ± 94 [3/4]	92 ± 28 [3/4]	28, 64 [1/2]
D	G 151 A	11 ± 4 [2/3]	6 ± 2 [2/3]	-
D	K 152 A	89, 73 [1/2]	46, 20 [1/2]	16, 35 [1/2]
3'D	L 153 A	95 [1/1]	83 [1/1]	46 [1/2]
3'D	K 154 A	175 ± 60 [3/4]	185 ± 84 [3/5]	82, 157 [1/2]
3'D	L 155 A	95, 128 [2/2]	115 [1/1]	40, 98 [2/2]
3'D	Y 156 A	105 [1/1]	65 ± 12 [3/3]	-
3'D	T 157 A	86, 104 [2/2]	89, 91 [2/2]	102, 104 [2/2]
3'D	G 158 A	84, 109 [2/2]	94, 89 [2/2]	77, 139 [2/2]
3'D	E 159 A	89, 97 [2/2]	101 ± 10 [2/3]	33, 96 [2/2]

FIG.34A

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C07K 14/505 | A3

(11) International Publication Number: WO 94/24160

(43) International Publication Date: 27 October 1994 (27.10.94)

(21) International Application Number: PCT/US94/04361

(22) International Filing Date: 21 April 1994 (21.04.94)

(30) Priority Data:
08/049,802 21 April 1993 (21.04.93) US

(71) Applicant: BRIGHAM AND WOMEN'S HOSPITAL [US/US]; 75 Francis Street, Boston, MA 02115 (US).

(72) Inventors: BOISSEL, Jean-Paul, R.; 15 Tanager Street, Arlington, MA 02174 (US). BUNN, H., Franklin; 184 Islington Road, Auburndale, MA 02166 (US). WEN, Danyi; 9 Sherman Road, Chestnut Hill, MA 02167 (US). SHOWERS, Mark, O.; 110 Longwood Avenue, Brookline, MA 02146 (US).

(74) Agents: CIMBALA, Michele, A. et al.; Sterne, Kessler, Goldstein & Fox, Suite 600, 1100 New York Avenue, N.W., Washington, DC 20005-3934 (US).

(81) Designated States: AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KG, KP, KR, KZ, LK, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
5 January 1995 (05.01.95)

(54) Title: ERYTHROPOIETIN MUTEINS WITH ENHANCED ACTIVITY

(57) Abstract

Novel modifications of erytropoietin (Epo) which improve the biological activity of the protein are provided. The modified Epo proteins (Epo muteins) may be used in the same manner as has been demonstrated for wild type Epo, except that relatively smaller doses are required due to the enhanced biological activity. Methods of using the Epo muteins for treatment of blood disorders are provided. A method of obtaining additional Epo muteins with enhanced biological activity is also provided.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

INTERNATIONAL SEARCH REPORT

Int onal Application No
PCT/US 94/04361

A. CLASSIFICATION F SUBJECT MATTER
IPC 5 C07K14/505

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 5 C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP,A,0 409 113 (BHERINGWERKE) 23 January 1991 cited in the application See the whole application, specifically Table 1	1-3,9, 13-25
Y	See the whole appliaction & AU,A,9 059 145 (BEHRINGWERKE) 19 July 1990	4-8, 10-12
X	EP,A,0 148 605 (KIRIN-AMGEN) 17 July 1985 cited in the application See page 21, lines 19-33, page 91, line5 to page 92, line 2, claims 1-13, 43, 35, 55-57 & US,A,4 703 008 (KIRIN-AMGEN) 27 October 1987	1,2, 13-20

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 November 1994	29. 11. 94
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Authorized officer Germinario, C

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP,A,0 357 804 (GENETICS INSTITUTE INC) 14 March 1990 cited in the application & US,A,4 835 260 (GENETICS INSTITUTEE) 30 May 1989 See the whole document	1-25
Y	EP,A,0 427 189 (SNOW BRAND MILK PRODUCTS & CO) 15 May 1991 See the whole document	1-25
A	WO,A,91 05867 (AMGEN) 2 May 1991 See claims33, 34	

1

INTERNATIONAL SEARCH REPORT

nformation on patent family members

Int onal Application No
PCT/US 94/04361

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-0409113	23-01-91	DE-A-	3923963	31-01-91
		AU-B-	638470	01-07-93
		AU-A-	5914590	24-01-91
		CA-A-	2021528	21-01-91
		JP-A-	3072885	28-03-91
AU-A-9059145		NONE		
EP-A-0148605	17-07-85	US-A-	4703008	27-10-87
		AU-A-	2042192	08-10-92
		AU-B-	600650	23-08-90
		AU-A-	3746785	26-06-85
		AU-B-	5272293	24-03-94
		AU-A-	5750490	04-10-90
		JP-A-	6093000	05-04-94
		JP-A-	1055190	02-03-89
		JP-B-	6055136	27-07-94
		JP-A-	3198792	29-08-91
		JP-B-	4035159	10-06-92
		JP-A-	3259098	19-11-91
		JP-B-	2017156	19-04-90
		JP-T-	61501627	07-08-86
		WO-A-	8502610	20-06-85
		CN-A-	1065019	07-10-92
		CN-A-	1062764	15-07-92
US-A-4703008	27-10-87	AU-A-	2042192	08-10-92
		AU-B-	600650	23-08-90
		AU-A-	3746785	26-06-85
		AU-B-	5272293	24-03-94
		AU-A-	5750490	04-10-90
		CN-A-	1065019	07-10-92
		CN-A-	1062764	15-07-92
		EP-A,B	0148605	17-07-85
		JP-A-	6093000	05-04-94
		JP-A-	1055190	02-03-89
		JP-B-	6055136	27-07-94
		JP-A-	3198792	29-08-91
		JP-B-	4035159	10-06-92
		JP-A-	3259098	19-11-91

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 94/04361

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US-A-4703008		WO-A- 8502610	20-06-85
		JP-B- 2017156	19-04-90
		JP-T- 61501627	07-08-86
EP-A-0357804	14-03-90	US-A- 4835260	30-05-89
US-A-4835260	30-05-89	EP-A- 0357804	14-03-90
EP-A-0427189	15-05-91	JP-A- 3151399	27-06-91
WO-A-9105867	02-05-91	AU-B- 646822	10-03-94
		AU-A- 6604290	16-05-91
		CA-A- 2027635	14-04-91
		CN-A- 1051936	05-06-91
		EP-A- 0428267	22-05-91
		JP-T- 4502331	23-04-92